

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 11:26:35 ; Search time 546 Seconds

(without alignments)  
2887.610 Million cell updates/sec

Title: US-10-023-597-24

Perfect score: 1577

Sequence: 1 MAENSSSVTEFLIAGLIHQ.....LRNKVXKALKRTPSRISFS 311

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3340653 segs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10023597/runat\_29092004\_163412\_14982/app\_query.fasta\_1.455  
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10023597 @CGN 1 1 480 @runat\_29092004\_163412\_14982  
-NCPU=6 -ICPU=3 -NO MMAP -LARGESUBERY -NEG\_SCORES=0 -WAIT -DSPELICK=100  
-LANGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Published Applications NA:  
1: /cgn2\_6/prodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*  
9: /cgn2\_6/prodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
10: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cgn2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/prodata/1/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/prodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgn2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgn2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result Query  
No. Match Length DB ID Description

1	1577	100.0	963	15	US-10-023-597-23	Sequence 23, Appl
2	1273.5	80.8	933	9	US-09-886-055-50	Sequence 50, Appl
3	1273.5	80.8	933	10	US-09-804-291-50	Sequence 50, Appl
4	1273.5	80.8	933	13	US-10-343-650A-395	Sequence 395, Appl
5	1273.5	80.8	933	15	US-10-220-382-35	Sequence 35, Appl
6	1273.5	80.8	1333	15	US-10-017-161-325	Sequence 325, Appl
7	1273.5	80.8	1333	16	US-10-292-798-291	Sequence 291, Appl
8	1268	80.4	956	16	US-10-024-399-1	Sequence 1, Appl
9	1266.5	80.3	953	16	US-10-005-041A-13	Sequence 13, Appl
10	1221	77.4	931	16	US-10-017-161-1301	Sequence 1301, Ap
11	1221	77.4	931	16	US-10-292-798-1069	Sequence 1069, Ap
12	1180	74.8	946	13	US-10-220-382-36	Sequence 36, Appl
13	1180	74.8	942	13	US-10-343-650A-397	Sequence 397, Appl
14	1180	74.8	951	16	US-10-387-625-47	Sequence 47, Appl
15	1180	74.8	1336	16	US-10-017-161-749	Sequence 749, Appl
16	1180	74.8	1336	16	US-10-292-798-653	Sequence 653, Appl
17	1058	67.1	952	15	US-10-023-597-59	Sequence 59, Appl
18	1048	66.5	961	15	US-10-023-601-89	Sequence 89, Appl
19	1024.5	65.0	934	15	US-10-023-597-125	Sequence 125, Appl
20	1023	64.9	981	15	US-10-023-597-81	Sequence 81, Appl
21	1013	64.2	943	15	US-10-023-597-57	Sequence 57, Appl
22	1006	63.8	950	15	US-10-023-597-87	Sequence 87, Appl
23	1006	63.8	953	15	US-10-023-597-63	Sequence 63, Appl
24	999.5	63.4	942	9	US-09-886-055-20	Sequence 20, Appl
25	999.5	63.4	942	10	US-09-804-291-20	Sequence 21, Appl
26	998.5	63.4	1342	15	US-10-017-161-321	Sequence 321, Appl
27	998.5	63.4	1342	16	US-10-292-798-287	Sequence 287, Appl
28	998.5	63.3	1342	15	US-10-017-161-315	Sequence 315, Appl
29	998.5	63.3	1342	16	US-10-292-798-281	Sequence 281, Appl
30	995.5	63.1	886	16	US-10-017-161-1259	Sequence 1259, Ap
31	995.5	63.1	886	16	US-10-292-798-1067	Sequence 1067, Ap
32	994.5	63.1	915	16	US-10-387-629-71	Sequence 71, Appl
33	992	62.9	981	15	US-10-023-597-91	Sequence 91, Appl
34	989	62.7	943	15	US-10-023-597-61	Sequence 61, Appl
35	978.5	62.0	942	16	US-10-387-629-73	Sequence 73, Appl
36	978.5	62.0	958	15	US-10-025-806-15	Sequence 15, Appl
37	978.5	62.0	974	15	US-10-024-399-35	Sequence 35, Appl
38	978.5	62.0	1342	15	US-10-017-161-313	Sequence 313, Appl
39	978.5	62.0	1342	16	US-10-292-798-279	Sequence 279, Appl
40	977	62.0	978	15	US-10-023-601-91	Sequence 91, Appl
41	977	62.0	1015	15	US-10-023-601-97	Sequence 97, Appl
42	976	61.9	942	15	US-10-023-601-99	Sequence 99, Appl
43	975	61.8	931	15	US-10-023-597-47	Sequence 47, Appl
44	972	61.6	943	15	US-10-023-597-89	Sequence 89, Appl
45	966.5	61.3	942	9	US-09-886-055-18	Sequence 18, Appl

## ALIGNMENTS

RESULT 1  
US-10-023-597-23  
; Sequence 23, Application US/10023597  
; Publication No. US20030109692A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Murallidhara  
; APPLICANT: Padigaru, Ramesh  
; APPLICANT: Li, Li  
; APPLICANT: Ballinger, Robert A.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Baugartner, Jason C.  
; APPLICANT: Buysse, Catherine E.  
; TITLE OF INVENTION: No. US20030109692A1 Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-224AD  
; CURRENT APPLICATION NUMBER: US/10/023, 597  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: 60/256, 635  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: 60/259, 743  
; PRIOR FILING DATE: 2001-01-04  
; PRIOR APPLICATION NUMBER: 60/299, 327  
; PRIOR FILING DATE: 2001-06-19

/ PRIOR APPLICATION NUMBER: 60/261,498  
 / PRIOR FILING DATE: 2001-01-12  
 / PRIOR APPLICATION NUMBER: 60/263,689  
 / PRIOR FILING DATE: 2001-01-24  
 / PRIOR APPLICATION NUMBER: 60/267,464  
 / PRIOR FILING DATE: 2001-02-08  
 / PRIOR APPLICATION NUMBER: 60/271,021  
 / PRIOR FILING DATE: 2001-02-22  
 / PRIOR APPLICATION NUMBER: 60/275,946  
 / PRIOR FILING DATE: 2001-03-14  
 / PRIOR APPLICATION NUMBER: 60/278,150  
 / PRIOR FILING DATE: 2001-03-23  
 / PRIOR APPLICATION NUMBER: 60/285,718  
 / PRIOR FILING DATE: 2001-04-03  
 / Remaining Prior Application data removed - See file Wrapper or PLM.  
 / NUMBER OF SEQ ID NOS: 128  
 / SOFTWARE: Patent Ver. 2.1  
 / SEQ ID NO: 23  
 / LENGTH: 963  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 US-10-023-597-23

Alignment Scores:  
 Pred. No.: 2,596-146 Length: 963  
 Score: 1577.00 Matches: 311  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0

US-10-023-597-24 (1-311) x US-10-023-597-23 (1-963)  
 QY 1 MetAlaAlaGluAsnSerSerSerValThrcGluPheIleuAlaGlyLeuIleHisGln 20  
 Db 2 ATGGGTGGCGAAGACTCTCCCTCCGTCAGAGTTATCCCTCGACGGCTTATCCACAG 61  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThValVal 40  
 Db 62 CCGGACCTCCAGTCCCGCT 121  
 QY 41 GlyAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArGLeuHisIleProMetTyr 60  
 Db 122 GGAACCTGGGCTTATATCTCTGATAGGCTCAACTCCGTCGACATATCCCATGTAC 181  
 QY PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProlys 80  
 Db 182 TTTTCCCTTCAACTGTCT 241  
 QY 81 MetLeuMetSerPheValSerArGlyAsnIleIleSerPheThrGlyCysMetSerGln 100  
 Db 242 ATGCTGATGAGTTTGTCTCAAGAGAAACATATTCTCTTCAAGGCTGTATGAGTCAG 301  
 QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120  
 Db 302 TTTTCT 361  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140  
 Db 362 GACGGCTACGCGGATCTGTACCACTGTGTACACGATCAACATGTCCTCCACAGTG 421  
 QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 Db 422 TGTGTGCTCTTAACTGCGGTCTACCGGATGGGGGTTTTTGGGGCTGTGGCTCATACA 481  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 Db 482 GGAATATAGTGTCTTCTACCTTTTGTGACACAACTTGTCAATCTCAATGTGTGAC 541  
 QY 181 IleuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 Db 542 ATCTTCCCTCTTGAAGTCTCTCTGCAACGGCTCTTACATTAATGTCTGTCATCTTT 601

QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 Db 602 ATTTGTGTACCGTTGGCATTTGGGAGCCCATTTGTCGGCTTTTATCTTATGTTT 661  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGlyValTyrSerIleValPheSerSer 240  
 Db 662 ATCTTTCAGCACTTCTCCGCTTATTCGTCGAGGCGAGGTCTAAACCTTCAGTAC 721  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyValAspThrTyrLeu 260  
 Db 722 TGCAGCTCTCACTAATTTGAGTTTCTTTCTTTGGTACGAGCTTTTACGTAACCTC 781  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280  
 Db 782 AAACCCCTTCCATTTTACCCCTCGACACAGGAGAAAGTCTCCCTGTTCTATACACT 841  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnValAspValIleLeuAla 300  
 Db 842 GTGGTCCCATGTTTATACCATTTATCTACACCTGAGGAAATAGAGATCAACTGCC 901  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 Db 902 CTGAGAGAACCTTTTCCAGAAATTAAGCTTTTCT 934

RESULT 2  
 US-09-886-055-50  
 / Sequence 50, Application US/09886055  
 / Patent No. US20020132273A1  
 / GENERAL INFORMATION:  
 / APPLICANT: STRYER, LUBERT  
 / APPLICANT: ZOZULYA, SERGEY  
 / TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
 / FILE REFERENCE: 078003-0277150  
 / CURRENT APPLICATION NUMBER: US/09/886,055  
 / PRIOR FILING DATE: 2001-06-22  
 / PRIOR APPLICATION NUMBER: 60/213,812  
 / NUMBER OF SEQ ID NOS: 522  
 / SOFTWARE: Patent Ver. 2.1  
 / SEQ ID NO: 50  
 / LENGTH: 933  
 / TYPE: DNA  
 / ORGANISM: Homo sapiens  
 US-09-886-055-50

Alignment Scores:  
 Pred. No.: 2,596-116 Length: 933  
 Score: 1273.50 Matches: 250  
 Percent Similarity: 87.46% Conservative: 22  
 Best Local Similarity: 80.39% Mismatches: 38  
 Query Match: 80.75% Indels: 1  
 DB: Gaps: 1

US-10-023-597-24 (1-311) x US-09-886-055-50 (1-933)  
 QY 1 MetAlaAlaGluAsnSerSerSerValThrcGluPheIleuAlaGlyLeuIleHisGln 20  
 Db 1 ATGGCAGCAAAAC---TCTTCTGTGACAGATTATCTCTGAAAGCTTAAACCAACAG 57  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThValVal 40  
 Db 58 CCGGACCTGGGATCCCT 117  
 QY 41 GlyAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArGLeuHisIleProMetTyr 60  
 Db 118 GGAACCTGGGCTTATTAACCTGATTTGGCTGAACCTTCACTGCAACATCCCATGTAC 177  
 QY PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProlys 80  
 Db 61 TTTTCTCTTTTAACTCTTTTAAATGATTTCTGTCTCTCACTACCATCACTCCCAA 237  
 QY 81 MetLeuMetSerPheValSerArGlyAsnIleIleSerPheThrGlyCysMetSerGln 100

```

Db      238 ATGCTATAGTGTTCCTCAAGAAAGACATCTTCCTTACAGGGGTATGACTCAG 297
Qy      101 PhepPhePheCysPhePheValPheSerGluSerPheIleuSerAlaMetValGlu 120
Db      298 CTTCTCTCTCTGCTCTTCTTTCGCTCTGAGTCTTCATCTCTGTCAGCGATGCGGTAT 357
Qy      121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThMetSerProGlnVal 140
Db      358 GACCGCTACGCGCATCTGTGAACCCACCTGTTGACACAGTCAACCTGCTTCCAGGTG 417
Qy      141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
Db      418 TGTTCCTCTCTTGTGTTGGTGCCTATGGGATGGGGTGTGCGGGCCATGGCCACACA 477
Qy      161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
Db      478 GGAAGCATATAGTAACCTGACCTTGTGCTGACCAACCTGTCAATCATATTCATGATGAC 537
Qy      181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
Db      538 ATCTTCCTCTCTCTTACGCTCTCTCTGCAACGCTCTTACATGATGAGCTGGGTCTTT 597
Qy      201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
Db      598 ATGTGTGCTGCTGTGACGTTGGAAATGCCATGTGCTGCTATTTCTTATGCGCTTC 657
Qy      221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerValAlaPheSerSer 240
Db      658 ATCTCTCCACGATTTTACACACAGATTTTACAGAGGAGGAGGCCAAACCTTTAGTACT 717
Qy      241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyValAlaPheThrTyrIleu 260
Db      718 TGCAGTTCACATATATGATGATTTCTTTCTTGTGTTCTGTTGCTTCAATGATATC 777
Qy      261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280
Db      778 AAACCCCTTTCATCTCCGCCCCGAGCAAGGAAAGTCCCTCCCTGTTCTATACATTA 837
Qy      281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnValAspValIleLeuAla 300
Db      838 ATAGTCCCTCTGTTAAACCATTAATCTATAGCTTACGAGACAGATGTCAAGATGCC 897
Qy      301 LeuLysArgThrPheSerArgIleSerPheSer 311
Db      898 CTGAGGAGAACTTGTGGCAGAAAAATCTTTCT 930

```

```

; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291-50

Alignment Scores:
Pred. No.: 2,59e-116 Length: 933
Score: 1273.50 Matches: 250
Percent Similarity: 87.46% Conservative: 22
Best Local Similarity: 80.39% Mismatches: 38
Query Match: 80.75% Indels: 1
DB: 10 Gaps: 1

US-10-023-597-24 (1-311) x US-09-804-291-50 (1-933)
Qy      1 MetAlaIaGluAsnSerSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20
Db      1 ATGGCAGCAAAAC--TCTTGTGACAGATTATCTCGAAGGCTTAACCCACAG 57
Qy      21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
Db      58 CCGGACGCGGAGCCCTCTCTCTCTGTTCTGAGTTTCTACAGCGTACCGGTG 117
Qy      41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
Db      118 GGGAACTCGGCTTGATTAACCTGATTTGGGTGAACTCAACCTGCAACATCCATCTAC 177
Qy      61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIleProLys 80
Db      178 TTCTTCTTTTAACTCTCTTATATGATTTCTGTTTCCATCTACATCACTCCCAA 237
Qy      81 MetLeuMetSerPheValSerArgValAsnIleIleSerPheThrGlyCysMetSerGln 100
Db      238 ATGCTATAGTGTTCCTCAAGAAAGACATCTTCCTTACAGGGGTATGACTCAG 297
Qy      101 PhePhePhePheCysPhePheValPheSerGluSerPheIleuSerAlaMetValGlu 120
Db      298 CTTCTCTCTCTGCTCTTCTTTCGCTCTGAGTCTTCATCTCTGTCAGCGATGCGGTAT 357
Qy      121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThMetSerProGlnVal 140
Db      358 GACCGCTACGCGCATCTGTGAACCCACCTGTTGACACAGTCAACCTGCTTCCAGGTG 417
Qy      141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
Db      418 TGTTCCTCTCTTGTGTTGGTGCCTATGGGATGGGGTGTGCGGGCCATGGCCACACA 477
Qy      161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
Db      478 GGAAGCATATAGTAACCTGACCTTGTGCTGACCAACCTGTCAATCATATTCATGATGAC 537
Qy      181 IleLeuProLeuLeuLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
Db      538 ATCTTCCTCTCTCTTACGCTCTCTCTGCAACGCTCTTACATGATGAGCTGGGTCTTT 597
Qy      201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
Db      598 ATGTGTGCTGCTGTGACGTTGGAAATGCCATGTGCTGCTATTTCTTATGCGCTTC 657
Qy      221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerValAlaPheSerSer 240
Db      658 ATCTCTCCACGATTTTACACACAGATTTTACAGAGGAGGAGGCCAAACCTTTAGTACT 717
Qy      241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyValAlaPheThrTyrIleu 260
Db      718 TGCAGTTCACATATATGATGATTTCTTTCTTGTGTTCTGTTGCTTCAATGATATC 777
Qy      261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280

```

```

Db      778 AAACCCCTTCCATCTGCGCCCTCGAGCAAGGAAAGTGTCTCCCTGTTTATATCCATA 837
Qy      281 ValValPromePheAsnProLeuIleTySerLeuArgAsnLysAspValLysLeuA 300
      838 ATAGTCCCGGTGTAAACCCCATTAATCTATAGCTGAGGAACAAGATGTCAAAGTTGCC 897
Qy      301 LeuLysArgThrPheSerArgLieserPheSer 311
      898 CTGAGGAGAACTTTGGGCGAGAAATACTTTCT 930

RESULT 4
US-10-343-650A-395
; Sequence 395, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
; APPLICANT: HAGA, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347
; CURRENT APPLICATION NUMBER: US/10/343, 650A
; PRIOR FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 395
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (1)..(933)
US-10-343-650A-395

Alignment Scores:
Pred. No.:      2,596-116      Length:      933
Score:          1273.50      Matches:      250
Percent Similarity: 87.46%      Conservative: 22
Best Local Similarity: 80.39%      Mismatches: 38
Query Match:    80.75%      Indels:      1
De:            13      Gaps:      1

US-10-023-597-24 (1-311) x US-10-343-650A-395 (1-933)
Qy      1 MetaAaAGLUsAsnSerSerValThrGluPheIleLeuAAGLyleuIleHISGLN 20
      1 ATGGCAGGCCAAAC---TCTTCTGTGACAGAGTTATCTCGAAGGTTAAACCAACAG 57
Qy      21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyraValValThValVal 40
      58 CCGGACCTGGGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 117
Qy      41 GlysLeuLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTy 60
      118 GGGAACTGGGCTGTATACCTGTATGCTGCTGAACCTTACCTGCACACCTCCATGTAC 177
Qy      61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrTrIleIleProLys 80
      178 TTTCTCTCTTTTAACTCTCTCTTTATAGATTCTGTTTCTCCACATCACTCCCAAA 237
Qy      81 MetLeuMetSerPheValSerArgLysAsnLieserPheThrGlyCysMetSerGln 100
      238 ATGCTGATGAGTTTGTCTCAAGAGAAACATCTTCTCTTCAAGAGGTATGACTCAG 297
Qy      101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerArgLysValGlu 120
      298 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 357
Qy      121 AspArgTyraValGlyIleCysAsnProLeuLeuTyrrIleTrMetSerProGlnVal 140
      358 GACCGCTAGTGCGCATCTGTAAACCACTGTGTACACAGTACCAATGTCTTGCCAGGTG 417

```

```

Qy      141 CysLeuLeuLeuLeuLeuGlyValTyrglyMetGlyValPheGlyAlaValAlaHisThr 160
      418 TGTGTGCTCTTTTGTGTGGTGCCTATGGATGGGTTTGCGGGCCATGGCCACACA 477
Qy      161 GlysAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrrMetCysAsp 180
      478 GAAAGCATATGAACTGACCTTCTGTGCTGACACCTTGTCAATATTTCAATGTGTGC 537
Qy      181 IleLeuProLeuLeuGlnLeuSerCysAsnGlySerTyrrIleAsnValLeuValIlePhe 200
      538 ATCTTCTCTCTCTTGAAGCTCTCTCTCAACAGCTCTTACATGATAGCTGTGGCTTT 597
Qy      201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrglyPhe 220
      598 ATGTGTGCTGTGTACCTGTGAATGCCATGTGACTGTCTTATTTTATGTGCTTC 657
Qy      221 IleLeuSerSerIleLeuArgValSerSerArgLysValArgSerLysAlaPheSerSer 240
      658 ATCTCTCCAGCATTTCTACACAAGTTCTTACAGAAGGAGGCTCCAAAGCTTTAGTACT 717
Qy      241 CysSerSerTyrrIleIleAlaValSerLeuPhePheGlySerGlyValPheThrTyrrLeu 260
      718 TGCAGTTCACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 777
Qy      261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrrThr 280
      778 AAACCCCTTCCATCTGCGCCCTCGAGCAAGGAAAGTGTCTCTCTCTGTATACCAATA 837
Qy      281 ValValPromePheAsnProLeuIleTySerLeuArgAsnLysAspValLysLeuA 300
      838 ATAGTCCCGGTGTAAACCCCATTAATCTATAGCTGAGGAACAAGATGTCAAAGTTGCC 897
Qy      301 LeuLysArgThrPheSerArgLieserPheSer 311
      898 CTGAGGAGAACTTTGGGCGAGAAATACTTTCT 930

*
RESULT 5
US-10-220-382-35
; Sequence 35, Application US/10220382
; Publication No. US2003011911A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LANL, Pireti
; APPLICANT: LANG, Y. Tom
; APPLICANT: PATTERSON, Chandza
; APPLICANT: YAO, Monique G.
; APPLICANT: SHIH, Leo L.
; APPLICANT: TRIBOUNEV, Catherine
; APPLICANT: LU, Dzung Aina M.
; APPLICANT: YUE, Henry
; APPLICANT: KHAN, Farrah A.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: YANG, Junning
; APPLICANT: HARLAND, Lee
; APPLICANT: WALSH, Roderick T.
; APPLICANT: LO, Terence P.
; APPLICANT: BOROMSKY, Mark L.
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0044 PCT
; CURRENT APPLICATION NUMBER: US/10/220,382
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature

```



OTHER INFORMATION: Incyte ID No. US20030119111A1 7472439CB1  
US-10-023-382-35

Alignment Scores:  
Pred. No.: 2,59e-116 Length: 933  
Score: 1273.50 Matches: 250  
Percent Similarity: 87.46% Conservative: 22  
Best Local Similarity: 80.39% Mismatches: 38  
Query Match: 80.75% Indels: 1  
Gaps: 1

US-10-023-597-24 (1-311) x US-10-023-382-35 (1-933)

```

QY 1 MetAlaAGluAsnSerSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20
Db 1 ATGGCAGGCCAATAAC---TCTTCTGTGACAGATTATCCCTGAAAGCTTAACCCACAG 57
QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyAlaValThrVal 40
Db 58 CCGGGACTGCGGATCCCTCTTCTCTGTTCTGCGTTCTTACACGGTCACCGTGTG 117
QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
Db 118 GGGAACTGGGCTTGATTAACCTGATTGGGCTGAACCTCACCCTGACACTCCCATGTAC 177
QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProlys 80
Db 178 TTCTTCTCTTTTAACTCTCTTAATAGATTTCGTGTTCTCCACTACATCACTCCCAAA 237
QY 81 MetLeuMetSerPheValSerArgValAsnIleIleSerPheThrGlyCysMetSerGln 100
Db 238 ATGCTGATAGTGTGTCTCAAGAAACATCATTTCTTACAGGGGTATGACTGAG 297
QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
Db 298 CTCTTCTTCTTCTGCTTCTTGTGCTGCTGATGCTCTCATCTGACGCGATGCGTAT 357
QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
Db 358 GACCGCTACGTGGCCATCTGTAAACCCACGTTGTACACGATCCATGCTTCCAGGG 417
QY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyValAlaValAlaHisThr 160
Db 418 TGTTCCTCTCTTCTTGTGGTCCCTATGGATGCGGTTGCGGGGCCATGCCCCACACA 477
QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
Db 478 GGAAGCATATGAACCTGACCTTCTGTGCTGACACCTTGTCAATCATTTCAATGTGTAC 537
QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
Db 538 ATCTTCTCTCTCTTGGCTCTCTGCTGACAGCTCTTACATGATAGCTGTGCTCTT 597
QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
Db 598 ATTGTGTGCTGTGATGCTGGAATGCAATGCTCATGCTCTTATTTCTATAGCCCTC 657
QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGlnGlyArgSerTyrAlaPheSerSer 240
Db 658 ATCTCTCCAGCATCTTACACACAGCTTACAGAAAGGAGGTCACAAAGCTTTATCTACT 717
QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
Db 718 TGCAGTCCCAATATGATTGTTCTCTTTGTTGTTCTGCTCTTTCATGATATC 777
QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280
Db 778 AAACCCCTTTCATCCCTGCGCCCTGACAAAGGAAATGTCCTCCGTTCTATACATCA 837
QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValIleLeuAla 300
Db 838 ATATGCTCCCGCTTAAACCCATTAATCTATAGCTTGAAGAAACAAGATGTCAAGATTGCC 897

```

QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
Db 898 CTGAGAGGAACATTGGGACAGAAAAATCTTTTCT 930

RESULT 6

```

US-10-017-161-325
; Sequence 325, Application US/100117161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OR INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 325
; LENGTH: 1333
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(1333)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1133)
; LOCATION: (201)..(1133)
US-10-017-161-325

```

Alignment Scores:  
Pred. No.: 4.32e-116 Length: 1333  
Score: 1273.50 Matches: 250  
Percent Similarity: 87.46% Conservative: 22  
Best Local Similarity: 80.39% Mismatches: 38  
Query Match: 80.75% Indels: 1  
Gaps: 1

US-10-023-597-24 (1-311) x US-10-017-161-325 (1-1333)

```

QY 1 MetAlaAGluAsnSerSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20
Db 201 ATGGCAGGCCAATAAC---TCTTCTGTGACAGATTATCCCTGAAAGCTTAACCCACAG 257
QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyAlaValThrVal 40
Db 258 CCGGACTGCGGATCCCTCTTCTCTGTTCTGCGTTCTTACACGGTCACCGTGTG 317
QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
Db 318 GGGAACTGGGCTTGATTAACCTGATTGGGCTGAACCTCACCCTGACACTCCCATGTAC 377
QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProlys 80
Db 378 TTCTTCTCTTTTAACTCTCTTAATAGATTTCGTGTTCTCCACTACATCACTCCAAA 437
QY 81 MetLeuMetSerPheValSerArgValAsnIleIleSerPheThrGlyCysMetSerGln 100
Db 438 ATGCTGATAGTGTGTCTCAAGAAACATCATTTCTTCTTACAGGGGTATGATCTGAG 497
QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
Db 498 CTCTTCTTCTTCTGCTTCTTGTGCTGTGAGTCTTCACTCATCGTACGCGATGCGTAT 557
QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
Db 558 GACCGCTACGTGGCCATCTGTAACCCACTGTGTACACAGTGTACACATGTCTTGCAGGTG 617
QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyValAlaValAlaHisThr 160

```

```

Db      618 TGTTCCTCCTTTGTTGGGCTTAATGGGATGGGCTTGTGGGCCATGGCCCAACA 677
Qy      161 GlysniIleValPheLeuThrPheCysAlaAspAsnIleuValAsnHisTyrMetCysAsp 180
      678 GGAAGCATATGAACTGACCTTCTGTGTCGACAACTGTCAATCATATTCATGTGTAC 737
Qy      181 IleuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIleuValIlePhe 200
      738 ATCTTCCTCTCTCTTASGCTCTCTCTGCAACAGCTCTTACATATGAGCTGGTGTCTT 797
Qy      201 IleValIleThrValIleGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
      798 ATTTGGTGGCTGTTCAGCTTGGAATGCCATTGTGCAGCTTTATTTCTTATTCCTCCTC 857
Qy      221 IleuSerSerIleLeuArgValSerSerAlaGluGlyArgSerIleValAlaPheSerSer 240
      858 ATCTCTCCAGCATTTCTACACAACAGTTCTACAGAGGAGGATCCAAAGCCCTTAGTACT 917
Qy      241 CysSerSerTyrIleIleAlaValSerLeuPheGlySerGlyAlaPheThrTyrIleu 260
      918 TGCAGTTCACATATATTTGACTTCTCTTTCTTTGTTGCTGTGCTTTCATGTATCTC 977
Qy      261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
      978 AAACCCCTTTCATCTGCTGCTCCCTCGACAGAGGAAGTGCTCTCTCTTCTATACCATTA 1037
Qy      281 ValValIleProMetPheAsnProLeuIleTyrSerIleuArgAsnLysAspValIleuVal 300
      1038 ATAGTCCCTCCGCTGTAAACCATTAATCTATAGCTTGAGGAACAAGATGTCAAAAGTTGCC 1097
Qy      301 LeuLysArgThrPheSerArgIleSerPheSer 311
      1098 CTGAGAGAGAACTTTGGGCAAAAATCTTTCT 1130

```

## RESULT 7

```

US-10-292-798-291
; Sequence 291, Application US/10292798
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: AIBURANT, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292, 798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 291
; LENGTH: 1133
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1333)
; NAME/KEY: CDS
; LOCATION: (201)..(1133)
US-10-292-798-291

```

## Alignment Scores:

```

Pred. No.: 4,32e-116 Length: 1333
Score: 1273.50 Matches: 250
Percent Similarity: 87.46% Conservative: 22
Best Local Similarity: 80.39% Mismatches: 38
Query Match: 80.75% Indels: 1
DB: 16 Gaps: 1

```

## US-10-023-597-24 (1-311) x US-10-292-798-291 (1-1333)

```

Qy      1 MetAlaIleGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
      201 ATGGCAACCAAAAC---TTTCTGTACAGACTTTATCTCGAAGGCTTAACCCACGAG 257
Qy      21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
      258 CCGGAGCTGGGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 317
Db      41 GlysniIleValPheLeuThrPheCysAlaAspAsnIleuValAsnHisTyrMetCysAsp 60
      318 GGAAGCATATGAACTGACCTTCTGTGTCGACAACTGTCAATCATATTCATGTGTAC 377
Qy      61 PhePheProPheAsnLeuSerLeuValAspPheSerPheThrThrIleIleProLys 80
      378 TTCTTCTCTTTTAACTCTCTTAAATAGATTTCTGTTTCTTCCATACCATCACTCCCAA 437
Db      81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
      438 ATGCTATAGATTGTCTCAAGAGAACATCATATTCCTTACAGGGGTGATGACTCAG 497
Qy      101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120
      498 CTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 557
Qy      121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
      558 GACCGCTACGTCGACATCTGTAAACCATCTGTGTGTACAGATCACCATGCTTGCAGAGTG 617
Qy      141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
      618 TGTTCCTCCTTTTGTGGTCCCTATGGAGTGGGATTTGCTGGGGCCATGGCCCAACA 677
Qy      161 GlysniIleValPheLeuThrPheCysAlaAspAsnIleuValAsnHisTyrMetCysAsp 180
      678 GGAAGCATATGAACTGACCTTCTGTGTCGACAACTGTCAATCATATTCATGTGTAC 737
Qy      181 IleuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIleuValIlePhe 200
      738 ATCTTCCTCTCTCTTASGCTCTCTCTGCAACAGCTCTTACATATGAGCTGGTGTCTT 797
Qy      201 IleValIleThrValIleGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
      798 ATTTGGTGGCTGTTCAGCTTGGAATGCCATTGTGCAGCTTTATTTCTTATTCCTCCTC 857
Qy      221 IleuSerSerIleLeuArgValSerSerAlaGluGlyArgSerIleValAlaPheSerSer 240
      858 ATCTCTCCAGCATTTCTACACAACAGTTCTACAGAGGAGGATCCAAAGCCCTTAGTACT 917
Qy      241 CysSerSerTyrIleIleAlaValSerLeuPheGlySerGlyAlaPheThrTyrIleu 260
      918 TGCAGTTCACATATATTTGACTTCTCTTTCTTTGTTGCTGTGCTTTCATGTATCTC 977
Qy      261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
      978 AAACCCCTTTCATCTGCTGCTCCCTCGACAGAGGAAGTGCTCTCTCTTCTATACCATTA 1037
Qy      281 ValValIleProMetPheAsnProLeuIleTyrSerIleuArgAsnLysAspValIleuVal 300
      1038 ATAGTCCCTCCGCTGTAAACCATTAATCTATAGCTTGAGGAACAAGATGTCAAAAGTTGCC 1097
Qy      301 LeuLysArgThrPheSerArgIleSerPheSer 311
      1098 CTGAGAGAGAACTTTGGGCAAAAATCTTTCT 1130

```

## RESULT 8

```

US-10-024-399-1
; Sequence 1, Application US/10024399
; Publication No. US20030100493A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara

```

APPLICANT: Kekuda, Ramesh  
 APPLICANT: Coleman, Steven D.  
 APPLICANT: Spytek, Kimberly A.  
 APPLICANT: Ballinger, Robert A.  
 APPLICANT: Vernet, Corine A.M.  
 APPLICANT: Li, Li  
 APPLICANT: Shenoy, Suresh G.  
 APPLICANT: Casman, Stacie J.  
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 FILE REFERENCE: 21402-224AE  
 CURRENT APPLICATION NUMBER: US/10/024,399  
 CURRENT FILING DATE: 2001-12-18  
 PRIOR APPLICATION NUMBER: 60/256,635  
 PRIOR FILING DATE: 2000-12-18  
 PRIOR APPLICATION NUMBER: 60/259,743  
 PRIOR FILING DATE: 2001-01-04  
 PRIOR APPLICATION NUMBER: 60/299,327  
 PRIOR FILING DATE: 2001-06-19  
 PRIOR APPLICATION NUMBER: 60/261,498  
 PRIOR FILING DATE: 2001-01-12  
 PRIOR APPLICATION NUMBER: 60/263,689  
 PRIOR FILING DATE: 2001-01-24  
 PRIOR APPLICATION NUMBER: 60/267,464  
 PRIOR FILING DATE: 2001-02-08  
 PRIOR APPLICATION NUMBER: 60/271,021  
 PRIOR FILING DATE: 2001-02-22  
 PRIOR APPLICATION NUMBER: 60/275,946  
 PRIOR FILING DATE: 2001-03-14  
 PRIOR APPLICATION NUMBER: 60/278,150  
 PRIOR FILING DATE: 2001-03-23  
 PRIOR APPLICATION NUMBER: 60/285,718  
 PRIOR FILING DATE: 2001-04-23  
 Remaining prior application data removed - See file wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 40  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 1  
 LENGTH: 976  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-10-024-399-1  
 Alignment Scores:  
 Pred. No.: 9,69e-116 Length: 976  
 Score: 1268.00 Matches: 247  
 Percent Similarity: 86.50% Conservative: 22  
 Best Local Similarity: 79.42% Mismatches: 42  
 Query Match: 80.41% Indels: 0  
 DB: 15 Gaps: 0  
 US-10-023-597-24 (1-311) x US-10-024-399-1 (1-976)  
 QY 1 MetAlaIaGluAnserserSerValThrGluPheIleuAlaGlyLeuIleHisGln 20  
 DB 8 ATGGGCCAGCCCAAACTCTTCGTGACAGATTATCTTCGAAGCTTAACCCACG 67  
 QY 21 ProGlyIeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrVal 40  
 DB 68 CCGGGAGCTGGGAGTCCCTCTTCTCTGTTCTGAGTTTCAACGGTCACCGTGATG 127  
 QY 41 GlyAsnIeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetI 60  
 DB 128 GGGAACTGGGCTTGATTAACCTGATTGGGTGAACCTCACTGCACACTCCCATGTAC 187  
 QY PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIlePro 80  
 DB 188 TTCCTCTTTTAACCTCTCTTAATAGATTTCGTTTCCATCACTACACTCCCAAA 247  
 QY 81 MetLeuMetSerPheValSerArgLeuAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 248 ATGCTGATGAGTTTGTCTCAAGAGAACAATCTTCTTCAACAGGCTGTATGACTCAG 307  
 QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

DB 308 CTCTCTCTTCTGCTCTTCTTGTGCTGCTGACCTCTTCATCTCTGACGATGGCGTAT 367  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuTyrThrIleThrMetSerProGlnVal 140  
 DB 368 GACGCTACGTCGACATCTGTATACCACTGTTTACAGTCCATGCTTCCAGGGG 427  
 QY 141 CysIleuLeuIeuIeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 DB 428 TGTGTCTCTTGTGTGTGCTGCTGATGGATGGGTTCTCTGGCCATGGCCACAC 487  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 DB 488 GGAAGCATATGAACTGACCTTCTGTGCTGACCAACTGTCAATCATTCATCATGTGAC 547  
 QY 181 IleLeuProIeuIeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 DB 548 ATCTTCTCTCTGCTGAGCTCTCTGCAACAGCTTTACATGATAGCTGTGTCTTT 607  
 QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 DB 608 ATTGTGTGCTGTGTGACGTGGAATGCCATTTGTCACTGTCTTATTTCTTATGCTTC 667  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerIleValPheSerSer 240  
 DB 668 ATCTCTCCAGCACTTTCACACACAGCTTTCACAGAAAGGCTCAAGCTTATGACT 727  
 QY 241 CysSerSerTyrIleIleAlaValIleSerLeuPheGlySerGlyAlaPheThrTyrLeu 260  
 DB 728 TGCAGTTCCACATATTTGATGTTCTCTTCTTCTGTGTCTTTCATGATGATCTC 787  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThr 280  
 DB 788 AAACCCCTTTCATCTGCGCCCTCGACCAAGGAATGTCCTCCCTGTTCTATACATA 847  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValIleLeuAla 300  
 DB 848 ATAGTCCCGGTGTAACCATTAATCTATGCTTGTGGAACAAAGATGTCAAAGTTGCC 907  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 DB 908 CTGAGAGAACTTTGGCGCAGAAAATCTTTCT 940  
 RESULT 9  
 US-10-005-041A-13  
 Sequence 13, Application US/10005041A  
 Publication No. US2003023231A1  
 GENERAL INFORMATION:  
 APPLICANT: Casman, Stacie J  
 APPLICANT: Padigar, Muralidhara  
 APPLICANT: Burgess, Catherine E  
 APPLICANT: Shinkens, Richard A  
 APPLICANT: Spytek, Kimberly A  
 APPLICANT: Gilbert, Jennifer A  
 APPLICANT: Mayotte, Jane E  
 APPLICANT: Baumgartner, Jason C  
 APPLICANT: Mishra, Vishnu  
 APPLICANT: Vernet, Corine AM  
 APPLICANT: Dickinson, Kevin S  
 APPLICANT: Ballinger, Robert A  
 APPLICANT: Wolenc, Adam R  
 APPLICANT: Edinger, Shlomit R  
 APPLICANT: MacDougall, John R  
 APPLICANT: Smithson, Glenda  
 APPLICANT: Ellerman, Karen  
 APPLICANT: Stone, David J  
 APPLICANT: Gunther, Erik  
 APPLICANT: Gerlach, Valerie  
 TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
 FILE REFERENCE: 21402-215  
 CURRENT APPLICATION NUMBER: US/10/005,041A  
 CURRENT FILING DATE: 2001-12-04  
 PRIOR APPLICATION NUMBER: 60/251,459  
 PRIOR FILING DATE: 2000-12-05

```

1  PRIOR APPLICATION NUMBER: 60/259,007
2  PRIOR FILING DATE: 2000-12-29
3  NUMBER OF SEQ ID NOS: 205
4  SOFTWARE: PatentIn Ver. 2.1
5  SEQ ID NO: 13
6  LENGTH: 953
7  TYPE: DNA
8  ORGANISM: Homo sapiens
9  IS-10-005-041a-13

```

Alignment Scores:	
Pred. No.:	1,32e-115
Score:	1266.50
Percent Similarity:	87.14%
Best local Similarity:	80.06%
Query Match:	80.31%
DB:	16
Gaps:	1
length:	953
Matches:	249
Conservative:	22
Mismatches:	39
Indels:	32
Gaps:	1

US-10-023-597-24 (1-311) X US-10-005-041A-13 (1-953)

QY	16	ATGGCAGCCAAAAC---TCTTCTGACAGAGTTTATCTCTGAGGCTTAAACCAACAG	72
QY	21	ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal	40
Db	73	CCGGAGCTGGGAGTCCCCCTCTCTCTCTGTTTCTGGGTTTCTAACAGGTCACCTGGG	133
QY	41	GlyAsnLeuGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	60
Db	133	GGAAACCTGGGCTTGAATACCTCGATGGGCTGAACTCCACCTCCACCTCCATGAC	192
QY	61	PhePheProPheAsnSerLeuValAspPheSerPheSerThrThrIleLeuProlys	80
Db	193	TTCCTCTTTTAACTCTCTTAAATAGATTCTGTTCTCCACTACATCACTCCAAA	253
QY	81	MetLeuMetSerPheValSerArglyValAsnIleIleSerPheThrGlyCysMetSerGln	100
Db	253	ATCTCATATAGTTTGTCTCAAGAAACATCATCTTCCTTCACAGGGTATGACTCAG	313
QY	101	PhePhePhePheCysPhePheValPheSerGlyUserPheIleLeuSerAlaMetValGlu	120
Db	313	CTTTCTTTCTTCCTGCTCTTGTCTGCTCTGAGTCTTCATCTGTGACGATGCGTAT	373
QY	121	AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal	140
Db	373	GACCGTAGTGGCCATCTGTAAACCACTGTTGTCACAGTCACCATCTCTTGGCAGTG	433
QY	141	CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr	160
Db	433	TGTTTGCTCTCTTGTGGTGCTCATATGGAGTGGGTTGCTGGGGCCATGGCCACACA	493
QY	161	GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp	180
Db	493	GGAAGCATATACACTGACCTCTCTGTCTGCAACCTTGTCATCATTTGATGTGAC	553
QY	181	IleLeuProLeuLeuGlnLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe	200
Db	553	ATCTTCTCTCTCTGAGCTCTCGGAACAGCTCTTACATGATGAGCCGGTGTCTT	613
QY	201	IleValIleValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe	220
Db	613	ATTGGGGTGGCTGTGAGCTGTGATGCCCATTTGCATCTTAAATTTCTTAAGCCCTC	673
QY	221	IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerIysAlaPheSerSer	240
Db	673	ATCTCTCCAGATTTCTACACACAGTTCTACAGAGGCAAGTCCAAAGCCTTTAGACT	733
QY	241	CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu	260
Db	733	TGCAGTCTCCACATATATGATGTTCTCTTTCTTGGTCTGTGCTTTCATGATATCTC	793
QY	261	LysProProSerIleLeuProLeuAspGlnGlyIysValSerSerLeuPheTyrThrThr	280

Db	793	AAACCCCTTTCATCCTGCCTCGACAGGAAAGTGTCTCTCCCTGTTCTATACATA	853
Oy	281	ValValPromecheasnpProleuileTyserLeuarganlysaapVallyLeuJa	300
Db	853	ATAGTCCCGGTGTAAACCATTAATCATATAGCTTGAGGAACAAGATGTCAAGTGGCC	912
Oy	301	LeuTyserThpseresergrileserheser	311
Db	913	CTAGAGGAACTTTGGCGAGAAAAATCTTTTCT	945

RESULT 10  
US-10-017

```

US-10-017-161-1301
? ; Sequence 1301: Application US/10017161
? Publication No. US20030143668A1
? GENERAL INFORMATION:
? APPLICANT: SOMA, MAKIRO
? APPLICANT: ASAI, KIYOSHI
? APPLICANT: AKIYAMA, YUTKA
? APPLICANT: ABURATANI, HIROYUKI
? TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
? FILE REFERENCE: 084335/0152
? CURRENT APPLICATION NUMBER: US/10/017,161
? CURRENT FILING DATE: 2002-12-18
? PRIOR APPLICATION NUMBER: JP 2001/246789
? PRIOR FILING DATE: 2001-06-18
? NUMBER OF SEQ ID NOS: 2430
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 1301
? LENGTH: 931
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: source
? LOCATION: (1)..(931)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (201)..(731)
? US-10-017-161-1301

```

Alignment Scores:	
Prad. No.:	4.04e-11
Score:	1.221.00
Percent Similarity:	99.59%
Best Local Similarity:	99.59%
Query Match:	77.43%
GB:	15
Length:	93
Matches:	24
Conservative:	0
Mismatch:	1
Indels:	0
Gaps:	0

US-10-023-597-24 (1-311) X US-10-017-161-1301 (1-931)

QY	70	AspPheSerPheSerThrThiIleIleProLysMetLeuMetSerPheValSerArgLys	89
Db	3	GAITTTAGTTCCTCAGACACATCATCTCCAAAATGCTGATGAGTTTGTCACGAAAG	62
QY	90	AsnIleIleSerPheThrGlyCysMetSerGlnPhePhePheCysPhePheValPhe	109
Db	63	AACATTATTTCTCCACACAGGCTGTATGATGATGATCTTCTCTCTCTGTTCTGTCTTT	122
QY	110	SerGlnSerPheIleIleLeuSerAlaMetValGluSparGlyArgValGlyIleCysAsnPro	129
QY	123	TCTGAGTCCTTCATCTCTCGCGCATGGTGTAGAGACCGCTACGTGGCACTGTACCCA	182
Db			
QY	130	LeuLeuTyrThrIleIleThrMetSerProGlnValCysLeuLeuLeuLeuGluValTyr	149
Db	183	CTGTGTACACGATCCACCATCTCTCCCCAGGTGTGTGCTCTTCTTACTAGGGGTGCTAC	242
QY	150	GlyMetGlyValPheGlyAlaAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCys	169
QY	243	GGATGGGGGGTTTTTGGGGGCTGTGGCTCCACACAGAAATTAATGTTTCTCACCTTTTGT	302
Db			
QY	170	AlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCys	189
QY	303	GCGACACACCTTGCAATCACTACATGTGTGAACATCTTCTCCCTTTTGAGCTCTCTGC	362



FILE REFERENCE: PI-0044 PCT  
 CURRENT APPLICATION NUMBER: US/10/220,382  
 CURRENT FILING DATE: 2001-03-01  
 PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730  
 PRIOR FILING DATE: 2000-03-03; 2000-03-10; 2000-03-17; 2000-03-20  
 NUMBER OF SEQ ID NOS: 42  
 SOFTWARE: PERL Program  
 SEQ ID NO 36  
 LENGTH: 936  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte ID No. US20030119111A1 7472440CBI  
 US-10-220-382-36

## Alignment Scores:

Pred. No.:	4,64e-107	Length:	936
Score:	1180.00	Matches:	226
Percent Similarity:	83.92%	Conservative:	35
Best Local Similarity:	72.67%	Mismatches:	50
Query Match:	74.83%	Indels:	0
DB:	15	Gaps:	0

US-10-023-597-24 (1-311) x US-10-220-382-36 (1-936)

QY 1 MetAlaAlaGluuSerSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20  
 Db 1 ATGCGTGTGAGAAATTCCTTCCTGACACAGTTATCTCGACGGCTTAAGTACGACCA 60  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheValValThrValVal 40  
 Db 61 CCGGAGTCCAGATCCCT 120  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 Db 121 GGGAACTGGGCTGTATTAACCTGATTAAGCTCACTCACTCAACACCCCTATGTAC 180  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIleProGly 80  
 Db 181 TTCTTCTCTATACCTGCT 240  
 QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100  
 Db 241 ATGCTATGAGCTTGTCTTAAAGAAACACATCTCTCTCGACGGGTATGACCTAG 300  
 QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120  
 Db 301 CTCTTCT 360  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlyVal 140  
 Db 361 GACCGCTATGTGCCATCTGTAAACCACTGTGTACATGACCATCTCTCCACAGTG 420  
 QY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 Db 421 TGTTCCTCTCTTGTGTGGGTCTCTATGGAGGGGTGTCTGGGGGCAATGGCCACACA 480  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 Db 481 GCGTGATGATGGGTGACCTTCTGTGCATTAACCTTGTCAACCATTAATGTGAC 540  
 QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValIleValIlePhe 200  
 Db 541 ATCTTCCCTCTCTGAGTGTCTGTGACACGACCATTAATGATGAGCTTGTATGTTT 600  
 QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 Db 601 GTTGTTGGGCAATTAATGTGTGTGTCACACAGCACACATCTTCTTCTTATGCTCTC 660  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240  
 Db 661 ATTCTCTCAGCATCTTCCACATGATTTACACGGAGGGCAGCTCCAAAGCTTACAGACC 720

QY 241 CysSerSerTyrIleIleAlaValSerLeuPheGlySerGlyValAlaPheThrTyrLeu 260  
 Db 721 TGGAGTCCACATATTAAGCAATTCCTGCTTGTGTGGTGTGAGACATTCAGTACTTC 780  
 QY 261 LysProProSerIleLeuProLeuAspGlnIlySerValSerSerLeuPheTyrThrThr 280  
 Db 781 AAACCTTCTCTCTTTTGTGCTATGACACAGGCAAGGTCTTCTCTATCTATACACT 840  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValIlyLeuAla 300  
 Db 841 GTGGTCCCTCTCTCAACCTCACTTAATTAATTAATTAATTAATTAATTAATTAATTA 900  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 Db 901 CTAAAGAAATCTTGACAAATAATGCAATTCCTCC 933

## RESULT 13

US-10-343-650A-397  
 Sequence 397, Application US/10343650A  
 Publication No. US20040067499A1  
 GENERAL INFORMATION:  
 APPLICANT: HAGA, TATSUYA  
 TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR  
 CURRENT APPLICATION NUMBER: US/10/343,650A  
 CURRENT FILING DATE: 2003-07-21  
 PRIOR APPLICATION NUMBER: JP 2000/237818  
 PRIOR FILING DATE: 2000-08-04  
 PRIOR APPLICATION NUMBER: JP 2001/34434  
 PRIOR FILING DATE: 2001-02-13  
 NUMBER OF SEQ ID NOS: 694  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 397  
 LENGTH: 942  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(942)  
 US-10-343-650A-397

Alignment Scores:  
 Pred. No.: 4,68e-107  
 Score: 1180.00  
 Percent Similarity: 83.92%  
 Best Local Similarity: 72.67%  
 Query Match: 74.83%

Pred. No.:	4,68e-107	Length:	942
Score:	1180.00	Matches:	226
Percent Similarity:	83.92%	Conservative:	35
Best Local Similarity:	72.67%	Mismatches:	50
Query Match:	74.83%	Indels:	0
DB:	13	Gaps:	0

US-10-023-597-24 (1-311) x US-10-343-650A-397 (1-942)

QY 1 MetAlaAlaGluuSerSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20  
 Db 7 ATGCGTGTGAGAAATTCCTTCCTGACACAGTTATCTCGACGGCTTAAGTACGACCA 66  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheValValThrValVal 40  
 Db 67 CCGGAGTCCAGATCCCT 126  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 Db 127 GGGAACTGGGCTGTATTAACCTGATTAAGCTCACTCACTTGACACCCCTATGTAC 186  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIleProGly 80  
 Db 187 TTCTTCTCTTAACTGTCT 246  
 QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100  
 Db 247 ATGCTATGAGCTTGTCTTAAAGAAACACATCTCTCTACACGAGGTATGACTCAG 306  
 QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120





APPLICANT: ASAI, KIYOSHI  
 APPLICANT: AKIYAMA, YUTAKA  
 APPLICANT: AUBRATANI, HIROYUKI  
 TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
 FILE REFERENCE: 084335/0152  
 CURRENT APPLICATION NUMBER: US/10/017,161  
 CURRENT FILING DATE: 2002-12-18  
 PRIOR APPLICATION NUMBER: JP 2001/246789  
 NUMBER OF SEQ ID NOS: 2430  
 PRIOR FILING DATE: 2001-06-18  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 749  
 LENGTH: 1336  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: source  
 LOCATION: (1)..(1336)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (201)..(1136)  
 US-10-017-161-749

Alignment Scores:  
 Pred. No.: 7,72e-107 Length: 1336  
 Score: 1180.00 Matches: 226  
 Percent Similarity: 83.92% Conservative: 35  
 Best Local Similarity: 72.67% Mismatches: 50  
 Query Match: 74.83% Indels: 0  
 DB: 15 Gaps: 0

US-10-023-597-24 (1-311) X US-10-017-161-749 (1-1336)

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20  
 DB 201 ATGGCTGCTGAGAAATTCCTCTCGTACAGATTATCTCGAGGCTTAACTGACCA 260  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValAl 40  
 DB 261 CCGGAGATCCAGATCCCT 320  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 DB 321 GGGAACTGGGCTTATACCTGATAGGCTCAACTCTCACTTGCACACCCCTATATGAC 380  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIleProLys 80  
 DB 381 TTCTTCTCTATACCTGCT 440  
 QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 441 ATCTGATGAGCTTTGTCTTAAAGAAAGACAGATCTCTACGCGAGGTATGACTCAG 500  
 QY 101 PhePhePhePheCysPhePheValPheSerGlnSerPheIleLeuSerAlaMetValGlu 120  
 DB 501 CTCTTCT 560  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140  
 DB 561 GACCGCTATGTGGCCATCTGTAAACCACTGTTGTACATGTCACCATGATGCTCCCAAG 620  
 QY 141 CysLeuLeuLeuLeuGlnValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 DB 621 TGTTCCT 680  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 DB 681 GCGTGATATGTGGGTGACCTTCTGTGCAATACCTGTTCAACCACTACATGTGTGAC 740  
 QY 181 IleLeuProLeuLeuGlnLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 DB 741 ATCTTCCCT 800

QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 DB 801 GTTGTTCGGGATGATATTGGTGTGCCACAGTACACATCTTCTCTATGCTCTC 860  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerIleAlaPheSerSer 240  
 DB 861 ATTCTCTCAGCATTTCCACATTGATCCAGAGGAGGTCCAAAGCCTTCAGCAC 920  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyValAlaPheThrTyrLeu 260  
 DB 921 TCCAGTCCACATATTGACAGTTCTCTGTTCTTGGGTGAGAGCATTCATCTACCTC 980  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280  
 DB 981 AAACCTTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1040  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300  
 DB 1041 GTGGTGCCCATGCTCAACCCATTATTTATAGCTGAGGAATAGAGCTCAAGTTGCT 1100  
 QY 301 LeuValArgThrPheSerArgIleSerPheSer 311  
 DB 1101 CTAAAGAAAATCTTGACAAATAATGCAATTCCTCC 1133

Search completed: September 30, 2004, 13:55:03  
 Job time : 553 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2004, 07:23:04 ; Search time 561 Seconds

(without alignments)  
8702.305 Million cell updates/sec

Title: US-10-023-597-23

Perfect score: 963

Sequence: 1 aatgctgcgcgaacacccct.....ttttagaacaagaaga 963

Scoring table: OLIGO\_NUC Gapex 60.0

Searched: 3340653 seqs, 2534783454 residues

Word size: 20

Total number of hits satisfying chosen parameters: 153

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCR\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCRUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	963	100.0	963	Sequence 23, App1
2	706	73.3	931	Sequence 1301, App1
3	706	73.3	931	Sequence 1069, App
4	121	12.6	886	Sequence 1067, App
5	121	12.6	886	Sequence 366, App
6	57	5.9	487	Sequence 359, App
7	33	4.0	487	Sequence 259, App
8	33	4.0	487	Sequence 348, App
9	33	4.0	487	Sequence 279, App
10	39	4.0	491	Sequence 313, App1
11	39	4.0	491	Sequence 50, App1
12	39	4.0	933	Sequence 395, App
13	39	4.0	933	Sequence 395, App
14	39	4.0	933	Sequence 395, App

15	39	4.0	933	US-10-220-382-35	Sequence 35, App1
16	39	4.0	933	US-10-005-004A-13	Sequence 13, App1
17	39	4.0	933	US-10-024-339-1	Sequence 1, App1
18	39	4.0	1333	US-10-017-161-325	Sequence 325, App
19	39	4.0	1333	US-10-292-798-291	Sequence 291, App
20	34	3.5	34	US-10-085-198-230	Sequence 230, App
21	33	3.4	485	US-09-747-155-232	Sequence 232, App
22	33	3.4	486	US-09-747-155-233	Sequence 233, App
23	33	3.4	936	US-10-220-382-36	Sequence 36, App1
24	33	3.4	932	US-10-343-650A-397	Sequence 397, App
25	33	3.4	931	US-10-387-629-47	Sequence 47, App1
26	33	3.4	1336	US-10-017-161-749	Sequence 749, App
27	33	3.4	1336	US-10-292-798-653	Sequence 653, App
28	30	3.1	945	US-10-300-846-13	Sequence 13, App1
29	28	2.9	945	US-10-023-597-95	Sequence 95, App1
30	28	2.9	950	US-10-023-597-113	Sequence 113, App
31	28	2.9	962	US-10-023-597-115	Sequence 115, App
32	26	2.7	951	US-09-886-055-134	Sequence 134, App
33	26	2.7	951	US-09-804-221-134	Sequence 134, App
34	26	2.7	951	US-10-343-650A-283	Sequence 283, App
35	26	2.7	951	US-10-387-629-21	Sequence 21, App1
36	26	2.7	971	US-10-433-581-21	Sequence 21, App1
37	26	2.7	980	US-10-025-806-49	Sequence 49, App1
38	26	2.7	980	US-10-025-806-51	Sequence 51, App1
39	26	2.7	1013	US-10-023-601-85	Sequence 85, App1
40	26	2.7	1331	US-10-017-161-921	Sequence 221, App
41	26	2.7	1331	US-10-017-161-931	Sequence 221, App
42	26	2.7	1351	US-10-292-798-191	Sequence 191, App
43	23	2.4	23	US-10-085-198-229	Sequence 229, App
44	23	2.4	507	US-09-777-789-45	Sequence 45, App1
45	23	2.4	528	US-09-777-789-40	Sequence 40, App1

## ALIGNMENTS

RESULT 1  
US-10-023-597-23  
Sequence 23, Application US/10023597  
Publication No. US20030109692A1  
GENERAL INFORMATION:  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Li, Li  
APPLICANT: Ballinger, Robert A.  
APPLICANT: Casman, Stacie J.  
APPLICANT: Soytek, Kimberly A.  
APPLICANT: Baumgartner, Jason C.  
APPLICANT: Burgess, Catherine E.  
TITLE OR INVENTION: No. US20030109692A1 Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-224AD  
CURRENT FILING DATE: US/10/023, 597  
PRIOR APPLICATION NUMBER: 2001-12-18  
PRIOR FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: 60/256, 635  
PRIOR FILING DATE: 2001-01-04  
PRIOR APPLICATION NUMBER: 60/259, 743  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: 60/299, 327  
PRIOR FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: 60/261, 498  
PRIOR FILING DATE: 2001-01-12  
PRIOR APPLICATION NUMBER: 60/263, 689  
PRIOR FILING DATE: 2001-01-24  
PRIOR APPLICATION NUMBER: 60/267, 464  
PRIOR FILING DATE: 2001-02-08  
PRIOR APPLICATION NUMBER: 60/271, 021  
PRIOR FILING DATE: 2001-02-22  
PRIOR APPLICATION NUMBER: 60/275, 946  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/278, 150  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: 60/285, 718  
PRIOR FILING DATE: 2001-04-03

Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 128  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 23  
 ; LENGTH: 963  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-023-597-23

Query Match 100.0%; Score 963; DB 15; Length 963;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 963; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATGGCTGCGGAGAACTCCCTCCGTCAGAGATTATCCCGAGGCTTAATCCCA 60  
 Db 1 AATGGCTGCGGAGAACTCCCTCCGTCAGAGATTATCCCGAGGCTTAATCCCA 60  
 QY 61 GCCGGAGCTCCAGATCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
 Db 61 GCCGGAGCTCCAGATCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
 QY 121 GGGGAACCTGGGCTTATATCTCTGA TAGGGCTCAACTCTGCGCTGCATATCCCATGTA 180  
 Db 121 GGGGAACCTGGGCTTATATCTCTGA TAGGGCTCAACTCTGCGCTGCATATCCCATGTA 180  
 QY 181 CTTTTCCTCCCTCACTTGTCCCTGATATTAGTTTCTCTACGACCAATCATTC 240  
 Db 181 CTTTTCCTCCCTCACTTGTCCCTGATATTAGTTTCTCTACGACCAATCATTC 240  
 QY 241 AATGCTGATGATTTGTCTCAAGGAAGAAATTAATTTCTTCAAGGGTGTATGAGCA 300  
 Db 241 AATGCTGATGATTTGTCTCAAGGAAGAAATTAATTTCTTCAAGGGTGTATGAGCA 300  
 QY 301 GTTCTTCT 360  
 Db 301 GTTCTTCT 360  
 QY 361 GGAACCGCTAGGAGGATCTGTAAACCACTGTTGTAAGATCAATCATTCGCCA 420  
 Db 361 GGAACCGCTAGGAGGATCTGTAAACCACTGTTGTAAGATCAATCATTCGCCA 420  
 QY 421 GTTCTTCT 480  
 Db 421 GTTCTTCT 480  
 QY 481 AGGAATATAGTGTCTCAACCTTTGTGACAGAAACCTTGCATCTCACTGATG 540  
 Db 481 AGGAATATAGTGTCTCAACCTTTGTGACAGAAACCTTGCATCTCACTGATG 540  
 QY 541 CATCTTCCCTCTCTGAGCTCTCTGCAAGGCTCTTAATAAATGCTGTGATCT 600  
 Db 541 CATCTTCCCTCTCTGAGCTCTCTGCAAGGCTCTTAATAAATGCTGTGATCT 600  
 QY 601 TATTGTGACCGTGGGATGGGGTGCCATGTGTGCGTTTATCTCTTANGGTT 660  
 Db 601 TATTGTGACCGTGGGATGGGGTGCCATGTGTGCGTTTATCTCTTANGGTT 660  
 QY 661 TATTGTGACCGTGGGATGGGGTGCCATGTGTGCGTTTATCTCTTANGGTT 720  
 Db 661 TATTGTGACCGTGGGATGGGGTGCCATGTGTGCGTTTATCTCTTANGGTT 720  
 QY 721 CTGACGCTCTCAATATGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
 Db 721 CTGACGCTCTCAATATGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 780  
 QY 781 CAAACCCCTTCATTTAACCCTGAGACGAGGAAAGTGTCTCTCTCTCTCTCT 840  
 Db 781 CAAACCCCTTCATTTAACCCTGAGACGAGGAAAGTGTCTCTCTCTCTCTCT 840  
 QY 841 TGTGAGGACCATGTTTAAACCATTAATCTACAGCTGAGAAATAGAGATGCA 900  
 Db 841 TGTGAGGACCATGTTTAAACCATTAATCTACAGCTGAGAAATAGAGATGCA 900

QY 901 CCTGAAGAGAACCTTTTCCAGATTAAGCTTTTCTTGAAAAAATTTAGAAACGAGAAA 960  
 Db 901 CCTGAAGAGAACCTTTTCCAGATTAAGCTTTTCTTGAAAAAATTTAGAAACGAGAAA 960  
 QY 961 AGA 963  
 Db 961 AGA 963

## RESULT 2

US-10-017-161-1301  
 ; Sequence 1301; Application US/10017161  
 ; Publication No. US2003014368A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SUMA, MAKIO  
 ; APPLICANT: ASAI, KIYOSHI  
 ; APPLICANT: AKIYAMA, YUTAKA  
 ; APPLICANT: ABURATANI, HIROYUKI  
 ; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
 ; FILE REFERENCE: 084335/0152  
 ; CURRENT APPLICATION NUMBER: US/10/017,161  
 ; PRIOR FILING DATE: 2002-12-18  
 ; PRIOR APPLICATION NUMBER: JP 2001/246789  
 ; NUMBER OF SEQ ID NOS: 2430  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1301  
 ; LENGTH: 931  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: source  
 ; LOCATION: (1)..(931)  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (201)..(731)  
 US-10-017-161-1301

Query Match 73.3%; Score 706; DB 15; Length 931;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 756; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 207 TAGATTAGTTAGTTCTCTACAGCATATCCCAAAATGCTGATGAGTTTGTCTCAAGA 266  
 Db 1 TAGATTAGTTAGTTCTCTACAGCATATCCCAAAATGCTGATGAGTTTGTCTCAAGA 266  
 QY 267 AGAATATATTCCTCTCAAGGGGTATAGTGTCTCTCTCTCTCTCTCTCTCT 326  
 Db 61 AGAATATATTCCTCTCAAGGGGTATAGTGTCTCTCTCTCTCTCTCTCTCT 326  
 QY 327 TTTCTGAGTCTCTTACCTCTGTGCGAGATGTGAGAGACGCTACGTTGCGATCTG 386  
 Db 121 TTTCTGAGTCTCTTACCTCTGTGCGAGATGTGAGAGACGCTACGTTGCGATCTG 386  
 QY 387 CACTGTGTACAGATCAACATGTCTCCCAAGGATGTGTGCTCTTACTGGGTGCT 446  
 Db 181 CACTGTGTACAGATCAACATGTCTCCCAAGGATGTGTGCTCTTACTGGGTGCT 446  
 QY 447 ACGGATGGGGGTTTTTGGGGCTGTGCTCATACAGAAATATAGTGTCTCACTTTT 506  
 Db 241 ACGGATGGGGGTTTTTGGGGCTGTGCTCATACAGAAATATAGTGTCTCACTTTT 506  
 QY 507 GTGACAGAACCTTGTCAATCACTACATGTGTGACATCTTCCCTTCTGAGCTCTCT 566  
 Db 301 GTGACAGAACCTTGTCAATCACTACATGTGTGACATCTTCCCTTCTGAGCTCTCT 566  
 QY 567 GCACGCTCTTCAATAAATGCTGTGATCTTATGTGTGACCGTTGGCATTTGGG 626  
 Db 361 GCACGCTCTTCAATAAATGCTGTGATCTTATGTGTGACCGTTGGCATTTGGG 626  
 QY 627 TGCCCATGTGTGCGGTTTTTATCTCTATAGTTTATCTTTCAGATCTCCGCGTTA 686  
 Db 421 TGCCCATGTGTGCGGTTTTTATCTCTATAGTTTATCTTTCAGATCTCCGCGTTA 686

Qy	68	GTTCTGTAGAGGACAGGTCTTAAGCCTCAGAGGTGAGCTCTTACATATTCAGATT	746
Db	481	GTTCTGTAGAGGACAGGTCTTAAGCCTCAGAGGTGAGCTCTTACATATTCAGATT	540
Qy	747	CTCTTTTCTTTGGGTACAGACTTTTACGTACCTTAACCCCTTCATTTTACCCCTGG	806
Db	541	CTCTTTTCTTTGGGTACAGAGCTTTTACGTACCTTAACCCCTTCATTTTACCCCTGG	600
Qy	807	ACCAAGGAGAAAGTGTCTCCCTGTCTTATACACTGTGTGTCACATGTTAACCCATTAA	866
Db	601	ACCAAGGAGAAAGTGTCTCCCTGTCTTATACACTGTGTGTCACATGTTAACCCATTAA	660
Qy	867	TCTACAGCCTGAGGAATANGATGTCAAACTTGCCTGTAGAGAACTTTTCCAGATTA	926
Db	661	TCTACAGCCTGAGGAATANGATGTCAAACTTGCCTGTAGAGAACTTTTCCAGATTA	720
Qy	927	GCTTTTCTGAAAAAATTTTGAACACGAAAAGAGA	963
Db	721	GCTTTTCTGAAAAAATTTTGAACACGAAAAGAGA	757

```

RESULT 3
US-10-292-798-1069
Sequence 1069. Application US/102922798
Publication No. US20030235833A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABRABATANI, HIROYUKI
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
FILE REFERENCE: 084335/166
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: 10/017,161
PRIOR FILING DATE: 2001-12-18
PRIOR APPLICATION NUMBER: JP 2001-246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2070
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1069
LENGTH: 931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: source
FEATURE:
LOCATION: (1)..(931)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(731)
US-10-292-798-1069

```

Query Match	73.3%	Score 706;	DB 16;	Length 931;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 756; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	207	TAGATTTAGATTCCTCAGACCATCAATCCCAAAATGCTATGATGATTTCTCCAGGA	266
Db	1	TGATTTAGATTCCTCAGACCATCAATCCCAAAATGCTATGATGATTTCTCCAGGA	60
QY	267	AGAACATTATTTCCCTCACAAGSGTGATGAGTCAGTCTCTCTCTCTGTTCTTTGCT	326
Db	61	AGAACATTATTTCCCTCACAAGSGTGATGAGTCAGTCTCTCTCTCTGTTCTTTGCT	120
QY	327	TTTCTAGTCCTTATCTCTGCGCATGCTGAGGACCGCTACGTGGGCACTCTGTAAC	386
Db	121	TTTCTAGTCCTTATCTCTGCGCATGCTGAGGACCGCTACCGTGGGCACTCTGTAAC	180
QY	387	CACGTGTGACGATACCATGTCCTCCAGTGTGTTGCTCTTTTAACTGGGTGCT	446
Db	181	CACGTGTGACGATACCATGTCCTCCAGTGTGTTGCTCTTTTAACTGGGTGCT	240

Qy	447	ACGGGATGGGGGTTTTTTGGGCGTGAGGCTCATPAAGGAAATATAGTGTTCACCTT	506
Dp	241	ACGGGATGGGGGTTTTTTGGGCGTGAGGCTCATPAAGGAAATATAGTGTTCACCTT	300
Qy	507	GTGCAGACAACTTGTGCATCATCATAGTGTGACATCTCTCCCTCTTGAAGCTCTCCT	566
Dp	301	GTGCAGACAACTTGTGCATCATCATAGTGTGACATCTCTCCCTCTTGAAGCTCTCCT	360
Qy	567	GCAGAGGCTCTACATTAATGTCCGTGCATCTTATGTGTGACCGTGGGATGGGG	626
Dp	361	GCAGAGGCTCTACATTAATGTCCGTGCATCTTATGTGTGACCGTGGGATGGGG	420
Qy	627	TGCCATTTGTSCCGTTTTTATCTCTTATGGTTTTATCTTTCAGCAITCTCCGGTTA	686
Dp	421	TGCCATTTGTSCCGTTTTTATCTCTTATGGTTTTATCTTTCAGCAITCTCCGGTTA	480
Qy	687	GTTCCTCTAGGGGAGGCTGTAAGGCTTCAGTACTGAGTCCCTACATATTCAGAGTT	746
Dp	481	GTTCCTCTAGGGGAGGCTGTAAGGCTTCAGTACTGAGTCCCTACATATTCAGAGTT	540
Qy	747	CTCTTTCTTGTGGTCAGAGCTTTTATGACTACTCAACCCCTCTTCATTTTACCCTGG	806
Dp	541	CTCTTTCTTGTGGTCAGAGCTTTTATGACTACTCAACCCCTCTTCATTTTACCCTGG	600
Qy	807	ACAGAGGGGAAAGTGCTCTCCCTGTTCTATACCACTGAGTGATCCCATGTTAAACCATTA	866
Dp	601	ACAGAGGGGAAAGTGCTCTCCCTGTTCTATACCACTGAGTGATCCCATGTTAAACCATTA	660
Qy	867	TCTACAGCCTGAGGAAATTAAGATGTCAAACTTGCCTTGAAGAGAACCTTTTCCAGATTA	926
Dp	661	TCTACAGCCTGAGGAAATTAAGATGTCAAACTTGCCTTGAAGAGAACCTTTTCCAGATTA	720
Qy	927	GCTTTCTTGAAAAAATTTTGAAGAACGAAAAAGGA	963
Dp	721	GCTTTCTTGAAAAAATTTTGAAGAACGAAAAAGGA	757

```

US-10-017-161-1299          RESULT 4
US-10-017-161-1299          /
Sequence 1299, Application US/10017161
Publication No. US20030143668A1
GENERAL INFORMATION:
APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABEHATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1299
LENGTH: 886
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: source
LOCATION: (1)..(886)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(686)
US-10-017-161-1299

```

Query Match	Score 121	DB 15	Length 886
Best Local Similarity	100.0%	Pred. NO. 8.1e-53	
Matches 121	Conservative 0	Indels 0	Gaps 0
QY	82	CTTCTCGGTTTCTAGGTTCTACGCATGACGGTGTGGGGAACCTGAGCTTATAAT	141



PRIOR FILING DATE: 1999-12-22  
NUMBER OF SEQ ID NOS: 431  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 299  
LENGTH: 487  
TYPE: DNA  
ORGANISM: Callithrix jacchus  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)..(487)  
OTHER INFORMATION: Taxon = 9483; gene = CJA171; Accession DDBJ/EMBL/GenBank = AF1798  
US-09-747-155-299  
OTHER INFORMATION: Product = olfactory receptor

Query Match 4.0%; Score 39; DB 9; Length 487;  
Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 TTTTATCTCTATAGGTTTATCTTTCAGCATTCCTCC 680  
DB 273 TTTTATCTCTATAGGTTTATCTTTCAGCATTCCTCC 311

RESULT 9  
US-09-747-155-348  
Sequence 348, Application US/09747155  
Patent No. US20020151692A1  
GENERAL INFORMATION:  
APPLICANT: Rouquier, Sylvie  
TITLE OF INVENTION: No. US20020151692A1 Polypeptides and Nucleic Acids Encoding Sam  
FILE REFERENCE: 19904-008 (C009B6834US)  
CURRENT APPLICATION NUMBER: US/09/747,155  
CURRENT FILING DATE: 2000-12-21  
PRIOR FILING DATE: 1999-12-22  
PRIOR APPLICATION NUMBER: 60/171,746  
NUMBER OF SEQ ID NOS: 431  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 348  
LENGTH: 487  
TYPE: DNA  
ORGANISM: Saimiri sciureus  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)..(487)  
OTHER INFORMATION: Taxon = 9521; gene = SSC195; Accession DDBJ/EMBL/GenBank = AF1798  
US-09-747-155-348  
OTHER INFORMATION: Product = olfactory receptor

Query Match 4.0%; Score 39; DB 9; Length 487;  
Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 TTTTATCTCTATAGGTTTATCTTTCAGCATTCCTCC 680  
DB 273 TTTTATCTCTATAGGTTTATCTTTCAGCATTCCTCC 311

RESULT 10  
US-09-747-155-279  
Sequence 279, Application US/09747155  
Patent No. US20020151692A1  
GENERAL INFORMATION:  
APPLICANT: Rouquier, Sylvie  
TITLE OF INVENTION: No. US20020151692A1 Polypeptides and Nucleic Acids Encoding Sam  
FILE REFERENCE: 19904-008 (C009B6834US)  
CURRENT APPLICATION NUMBER: US/09/747,155  
CURRENT FILING DATE: 2000-12-21

PRIOR APPLICATION NUMBER: 60/171,746  
PRIOR FILING DATE: 1999-12-22  
NUMBER OF SEQ ID NOS: 431  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 279  
LENGTH: 489  
TYPE: DNA  
ORGANISM: Macaca sylvanus  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)..(489)  
OTHER INFORMATION: Taxon = 9546; gene = MSY173; pseudogene; Accession DDBJ/EMBL/GenB.  
US-09-747-155-279

Query Match 4.0%; Score 39; DB 9; Length 489;  
Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 TTTTATCTCTATAGGTTTATCTTTCAGCATTCCTCC 680  
DB 273 TTTTATCTCTATAGGTTTATCTTTCAGCATTCCTCC 311

RESULT 11  
US-09-747-155-313  
Sequence 313, Application US/09747155  
Patent No. US20020151692A1  
GENERAL INFORMATION:  
APPLICANT: Rouquier, Sylvie  
TITLE OF INVENTION: No. US20020151692A1 Polypeptides and Nucleic Acids Encoding Sam  
FILE REFERENCE: 19904-008 (C009B6834US)  
CURRENT APPLICATION NUMBER: US/09/747,155  
CURRENT FILING DATE: 2000-12-21  
PRIOR FILING DATE: 1999-12-22  
PRIOR APPLICATION NUMBER: 60/171,746  
NUMBER OF SEQ ID NOS: 431  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 313  
LENGTH: 491  
TYPE: DNA  
ORGANISM: Pongo pygmaeus  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)..(491)  
OTHER INFORMATION: Taxon = 9600; gene = PLY110; pseudogene; Accession DDBJ/EMBL/GenB  
US-09-747-155-313

Query Match 4.0%; Score 39; DB 9; Length 491;  
Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 642 TTTTATCTCTATAGGTTTATCTTTCAGCATTCCTCC 680  
DB 273 TTTTATCTCTATAGGTTTATCTTTCAGCATTCCTCC 311

RESULT 12  
US-09-886-055-50  
Sequence 50, Application US/09886055  
Patent No. US20020132273A1  
GENERAL INFORMATION:  
APPLICANT: STRYER, LOBERT  
APPLICANT: ZOZULYA, SERGEY  
TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND  
FILE REFERENCE: 078003-0277150  
CURRENT APPLICATION NUMBER: US/09/886,055  
CURRENT FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 60/213,812  
PRIOR FILING DATE: 2000-06-22  
NUMBER OF SEQ ID NOS: 522  
SOFTWARE: PatentIn Ver. 2.1

```

; SEQ ID NO 50
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-055-50

```

```

Query Match          4.0%; Score 39; DB 9; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      235 TCCCAAAATGCTGATGAGTTTGTCTCAAGAGAGAACAT 273
DB      231 TCCCAAAATGCTGATGAGTTTGTCTCAAGAGAGAACAT 269

```

```

RESULT 13
US-09-804-291-50

```

```

; Sequence 50, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:

```

```

; APPLICANT: ZOZUIA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0278005
; CURRENT APPLICATION NUMBER: US/09/804,291
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914
; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291-50

```

```

Query Match          4.0%; Score 39; DB 10; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      235 TCCCAAAATGCTGATGAGTTTGTCTCAAGAGAGAACAT 273
DB      231 TCCCAAAATGCTGATGAGTTTGTCTCAAGAGAGAACAT 269

```

```

RESULT 14
US-10-343-650A-395

```

```

; Sequence 395, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
; APPLICANT: HAGA, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347
; CURRENT APPLICATION NUMBER: US/10/343,650A
; PRIOR FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434

```

```

; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 395
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(933)
US-10-343-650A-395

```

```

Query Match          4.0%; Score 39; DB 13; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      235 TCCCAAAATGCTGATGAGTTTGTCTCAAGAGAGAACAT 273
DB      231 TCCCAAAATGCTGATGAGTTTGTCTCAAGAGAGAACAT 269

```

```

RESULT 15
US-10-220-382-35

```

```

; Sequence 35, Application US/10220382
; Publication No. US20030119111A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: PATTERSON, Chandra
; APPLICANT: YAO, Monique G.
; APPLICANT: SHIH, Leo L.
; APPLICANT: TRIBOUTLEY, Catherine
; APPLICANT: LU, Dying, Aina M.
; APPLICANT: YUE, Henry
; APPLICANT: KHAN, Farrah A.
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: YANG, Junming
; APPLICANT: HARLAND, Lee
; APPLICANT: WALSH, Roderick T.
; APPLICANT: LO, Terence P.
; APPLICANT: BOROMSKY, Mark L.
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0044 PCT
; CURRENT APPLICATION NUMBER: US/10/220,382
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730
; PRIOR FILING DATE: 2000-03-03; 2000-03-10; 2000-03-17; 2000-03-20
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 35
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030119111A1 7472439CBI
US-10-220-382-35

```

```

Query Match          4.0%; Score 39; DB 15; Length 933;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      235 TCCCAAAATGCTGATGAGTTTGTCTCAAGAGAGAACAT 273
DB      231 TCCCAAAATGCTGATGAGTTTGTCTCAAGAGAGAACAT 269

```

Search completed: September 30, 2004, 10:01:34  
 J00 time : 562 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 15:18:29 | Search time 546 Seconds

(without alignment) 2887.610 Million cell updates/sec

Title: US-10-023-597-24

Perfect score: 311

Sequence: 1 MAENSSSVTEFLAGLIHQ.....LRNKDVKLXRTSRISFS 311

Scoring table:

OLIGO  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3340653 seqs, 2534763454 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6672343

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-O=/cgm2\_1/USPTO.spool/US10023597/rnat.29092004.163433.15514/app.query.fasta\_1.455  
-DB=Published Applications NA -OEXT=fastap -SUFFIX=olig.rnpb -MINMATCH=0.1  
-LOOPCU=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=oligo  
-TRANS-human40.csi -LIST=45 -DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1  
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=50  
-MAXLEN=200000000 -USER=US10023597 @CGM 1 1 480 @rnat.29092004.163433.15514  
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60  
-FGAPOF=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cgm2\_6/prodata/1/pubpna/US07\_FU0COMB.seq:\*  
2: /cgm2\_6/prodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgm2\_6/prodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgm2\_6/prodata/1/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgm2\_6/prodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgm2\_6/prodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgm2\_6/prodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgm2\_6/prodata/1/pubpna/US09\_PUBCOMB.seq:\*  
9: /cgm2\_6/prodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgm2\_6/prodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
11: /cgm2\_6/prodata/1/pubpna/US09C\_NEW\_PUB.seq:\*  
12: /cgm2\_6/prodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgm2\_6/prodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgm2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
15: /cgm2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgm2\_6/prodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
17: /cgm2\_6/prodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgm2\_6/prodata/1/pubpna/US60\_NEW\_PUB.seq:\*  
19: /cgm2\_6/prodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

1	311	100.0	963	15	US-10-023-597-23	Sequence 23, Appl
2	191	61.4	931	16	US-10-017-161-1301	Sequence 1301, Ap
3	191	61.4	931	16	US-10-292-798-1069	Sequence 1069, Ap
4	40	12.9	886	15	US-10-017-161-1299	Sequence 1299, Ap
5	40	12.9	886	16	US-10-292-798-1067	Sequence 1067, Ap
6	33	10.6	487	9	US-09-747-155-366	Sequence 366, App
7	20	6.4	485	9	US-09-747-155-232	Sequence 232, App
8	20	6.4	485	9	US-09-886-055-50	Sequence 50, Appl
9	20	6.4	933	13	US-09-804-291-50	Sequence 395, Appl
10	20	6.4	933	13	US-10-343-650A-395	Sequence 13, Appl
11	20	6.4	933	16	US-10-005-041A-13	Sequence 1, Appl
12	20	6.4	976	15	US-10-024-399-1	Sequence 325, App
13	20	6.4	1333	15	US-10-017-161-325	Sequence 291, App
14	20	6.4	1333	16	US-10-292-798-291	Sequence 359, App
15	19	6.1	486	9	US-09-747-155-359	Sequence 348, App
16	19	6.1	487	9	US-09-747-155-348	Sequence 67, Appl
17	19	6.1	956	15	US-10-023-597-67	Sequence 93, Appl
18	19	6.1	993	15	US-10-023-601-93	Sequence 289, App
19	19	5.8	487	9	US-09-747-155-299	Sequence 97, Appl
20	18	5.8	1015	15	US-10-023-601-97	Sequence 233, App
21	18	5.8	486	9	US-09-747-155-233	Sequence 279, App
22	17	5.5	489	9	US-09-747-155-279	Sequence 1383, Ap
23	17	5.5	871	15	US-10-017-161-1383	Sequence 1129, Ap
24	17	5.5	871	15	US-10-292-798-1129	Sequence 125, App
25	17	5.5	871	16	US-10-025-806-13	Sequence 63, Appl
26	17	5.5	934	15	US-10-023-597-125	Sequence 3264, Ap
27	16	5.1	953	15	US-10-023-597-63	Sequence 431, App
28	15	4.8	554	15	US-10-029-386-3264	Sequence 251, App
29	15	4.8	801	13	US-10-343-650A-431	Sequence 36, Appl
30	15	4.8	933	16	US-10-367-629-251	Sequence 397, App
31	15	4.8	936	15	US-10-220-382-36	Sequence 127, App
32	15	4.8	942	13	US-10-343-650A-397	Sequence 47, Appl
33	15	4.8	945	15	US-10-023-597-127	Sequence 9, Appl
34	15	4.8	951	16	US-10-387-629-47	Sequence 11, Appl
35	15	4.8	955	16	US-10-024-212-47	Sequence 13, Appl
36	15	4.8	955	16	US-10-024-212-11	Sequence 15, Appl
37	15	4.8	955	16	US-10-024-212-13	Sequence 35, Appl
38	15	4.8	955	16	US-10-024-212-15	Sequence 45, App
39	15	4.8	955	16	US-10-024-212-35	Sequence 47, App
40	15	4.8	974	15	US-10-017-161-545	Sequence 36, Appl
41	15	4.8	1333	15	US-10-292-798-477	Sequence 11, Appl
42	15	4.8	1333	16	US-10-182-822A-36	
43	15	4.8	1334	13	US-10-219-834-11	
44	15	4.8	1334	13		
45	15	4.8	1334	15		

## ALIGNMENTS

RESULT 1  
US-10-023-597-23  
Sequence 23, Application US/10023597  
Publication No. US20030109692A1  
GENERAL INFORMATION:  
APPLICANT: Padigar, Muralidhara  
APPLICANT: Kexuda, Ramesh  
APPLICANT: Li, Li  
APPLICANT: Ballinger, Robert A.  
APPLICANT: Casman, Steacie J.  
APPLICANT: Spytek, Kimberly A.  
APPLICANT: Baumgartner, Jason C.  
APPLICANT: Burgess, Catherine E.  
TITLE OF INVENTION: No. US20030109692A1e1 Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-224AD  
CURRENT APPLICATION NUMBER: US/10/023,597  
CURRENT FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: 60/256,635  
PRIOR FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: 60/259,743  
PRIOR FILING DATE: 2001-01-04  
PRIOR APPLICATION NUMBER: 60/299,327  
PRIOR FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: 60/261,498  
 PRIOR FILING DATE: 2001-01-12  
 PRIOR APPLICATION NUMBER: 60/263,689  
 PRIOR FILING DATE: 2001-01-24  
 PRIOR APPLICATION NUMBER: 60/267,464  
 PRIOR FILING DATE: 2001-02-08  
 PRIOR APPLICATION NUMBER: 60/271,021  
 PRIOR FILING DATE: 2001-02-22  
 PRIOR APPLICATION NUMBER: 60/275,946  
 PRIOR FILING DATE: 2001-03-14  
 PRIOR APPLICATION NUMBER: 60/278,150  
 PRIOR FILING DATE: 2001-03-23  
 PRIOR APPLICATION NUMBER: 60/285,718  
 PRIOR FILING DATE: 2001-04-03  
 Remaining Prior Application data removed - See file wrapper or PALM.  
 NUMBER OF SEQ ID NOS: 128  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 23  
 LENGTH: 963  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-10-023-597-23

Alignment Scores:  
 Pred. No.: 5,24e-300 Length: 963  
 Score: 311.00 Matches: 311  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 15 Gaps: 0

US-10-023-597-24 (1-311) x US-10-023-597-23 (1-963)

QY 1 MetaLaalagLusSerSerSerValThrgluPheileLeuAlaGlyLeuileHieglr 20  
 Db 2 AAGGCTGGCGAAGTCTCCCTCCGTCAGAGGTTATCCTCGAGGCTTATCCACCG 61  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyraIValThVal 40  
 Db 62 CCGGAGCTCCAGGTCGCCGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121  
 QY 41 GlyaenLeuGlyLeuile 60  
 Db 122 GGAACCTCGGCTTATATCTCTGATAGGCTCAATCTCGCTGCAATATCCCATGAC 181  
 QY 61 PhePheProPheAsnLeuSerleuValAspPheSerPheSerThrThrileProlys 80  
 Db 182 TTTTCCCTTCAACTGCTCCCTCGTAGATTGTTCTCTCTCTCTCTCTCTCTCTCTCT 241  
 QY 81 MetLeuMetSerPheValSerArgLysAsnilelelelelelelelelelelelelelele 100  
 Db 242 ATGCTGATGAGTTTGTCTCAAGAGAACATTAATCTCTCTCTCTCTCTCTCTCTCTCT 301  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheilelelelelelelelelele 120  
 Db 302 TTTCT 361  
 QY 121 AspArgTyraIleGlyIleCysAsnProleuLeuTyThrileThrMetSerProGlnVal 140  
 Db 362 GACCGCTACGAGGAGATCTGTAACCACTGTTGTACACGATCACAGGCTCTCCCAAGGTG 421  
 QY 141 CysLeuLeuLeuLeuLeuGlyValTyGlyMetGlyValPheGlyValValAlaHieThr 160  
 Db 422 TGTGTGCT 481  
 QY 161 Glyaenile 180  
 Db 482 GGAATATAGGTTCTCTCAACCTTGTGACAGCAACCTTGTCAATCATCATGTGTGAC 541  
 QY 181 IleLeuProleuLeuGlyLeuSerCysAsnGlySerTyThrilelelelelelelelele 200  
 Db 542 ATCTTCCCTCTCTCTGAGCTCTCTCTGCAAGGCTCTCTCTCTCTCTCTCTCTCTCT 601

QY 201 IleValAlaThrValGlyIleGlyValProileValAlaValPheilelelelelelelele 220  
 Db 602 ATTGTGTGACCCGTGTCATGAGGTCGATGAGGTCATGAGGTCATGAGGTCATGAGGTC 661  
 QY 221 IleLeuSerSerile 240  
 Db 662 ATTCTTCCAGCATCTCCGCGTGTGCTGAGGCGGCTTAAGGCTTCAAGCTTCAAGCTT 721  
 QY 241 CysSerSerTyThrile 260  
 Db 722 TGCAGCTCTACATTAATGTCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 781  
 QY 261 LysProProSerile 280  
 Db 782 AAACCCCTTCCATTTTACCCCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841  
 QY 281 ValValProMetPheAsnProleuile 300  
 Db 842 GTGCTGCCATGTTTAAACCATTAATCTACAGCTCTAGGATATGATGATGATGATGATG 901  
 QY 301 LeuLeuArgThrPheSerArgIleSerPheSer 311  
 Db 902 CTGAAGAGAACCTTTCCAGAAATTAAGCTTTTCT 934

# RESULT 2

US-10-017-161-1301  
 Sequence 1301, Application US/10017161  
 Publication No. US20030143668A1  
 GENERAL INFORMATION:  
 APPLICANT: SUMA, MAKIKO  
 APPLICANT: ASAI, KIYOSHI  
 APPLICANT: AKIYAMA, YUTAKA  
 APPLICANT: ABEURATANI, HIROYUKI  
 TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
 FILE REFERENCE: 084335/0152  
 CURRENT FILING DATE: 2002-12-18  
 PRIOR APPLICATION NUMBER: US/10/017,161  
 PRIOR FILING DATE: 2001-06-18  
 NUMBER OF SEQ ID NOS: 2430  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 1301  
 LENGTH: 931  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: source  
 LOCATION: (1)..(931)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (201)..(731)  
 US-10-017-161-1301

Alignment Scores:  
 Pred. No.: 2,15e-160 Length: 931  
 Score: 191.00 Matches: 241  
 Percent Similarity: 99.18% Conservative: 0  
 Best Local Similarity: 99.18% Mismatches: 1  
 Query Match: 61.41% Indels: 2  
 DB: 15 Gaps: 0

US-10-023-597-24 (1-311) x US-10-017-161-1301 (1-931)

QY 70 AspPheSerPheSerThrThrile 89  
 Db 3 GATTTAGTTCTCTACACCATCATTCCTCAAAATGCTGATGATGATGATGATGATGATG 62  
 QY 90 Asnile 109  
 Db 63 AACATTTCTCTCTCAAGGAGTATGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 122  
 QY 110 SerGlySerPheile 129

```

Db      123 TGTGAGCTCTTCATCTCTGCGGCGATGATGTA-GGACCGCTACGNGGCGATCTGTAAACC 181
Qy      129 OLeuLeuYrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuGlyValTy 149
Db      182 ACTGTTGTAACAGATACCAATGTCCTCCAGGTGTTGCTCTTACTGAGGTGCTA 241
Qy      149 rGlyMetGlyValPheGlyAlaValAlaIshisThrGlyAsnIleValPheLeuThrPheCy 169
Db      242 CCGGATGGGGGCTTTTGGGGGCTGTGGCTCATACAGAAATATAGTGTTCACCTTTTG 301
Qy      169 sAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGlyLeuSerCy 189
Db      302 TGCAGACAACTTGCATCACTACATGATGACATCTCCCTCCCTTGTGAGCTCTCTG 361
Qy      189 sAsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyVa 209
Db      362 CAACGGCTCTTACATAATGTCCTGCTCATCTTTATTTGTTGACCGCTTGGCATTGGGGT 421
Qy      209 lProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSe 229
Db      422 GCCCATGTTGCGCTTTTATCTTATGCTTTTATCTTTCACACATCTCCCGCGTTAG 481
Qy      229 rSerIleGlyValArgSerIleValPheSerSerCysSerSerTyrIleIleAlaValSe 249
Db      482 TTCTGCTGAGGCGAGGTCATAAGCCTTCAGTACCTGCACTCTTCATATATTGCAATTTC 541
Qy      249 rLeuPhePheGlySerGlyValaPheThrTyrLeuIleLeuSerSerIleLeuArgValSe 269
Db      542 TCTTTTCTTTGGGTGAGGAGCTTTTACGTAACCTCAACCCCTTCATTTTACCCCTGGA 601
Qy      269 pGlnGlyIleValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuI 289
Db      602 CCAAGGGAAGTGTCTCTCCCTGTTATACCACTGAGGCGCCATGTTTAAACCATTTAT 661
Qy      289 eTyrSerLeuArgAsnIleAspValIleLeuValIlePheIleValThrPheSerArgIleSe 309
Db      662 CTACAGCCTTGAGGAAATAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGATTAAG 721
Qy      309 rPheSer 311
Db      722 CTTTCT 728

RESULT 3
US-10-292-798-1069
; Sequence 1069, Application US/10292798
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292, 798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1069
; LENGTH: 931
; TYPE: DNA
; ORGANISM: Homo sapiens
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(931)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(731)
US-10-292-798-1069

```

```

Alignment Scores:
Pred. No.: 2,15e-180 Length: 931
Score: 191.00 Matches: 241
Percent Similarity: 99.18% Conservative: 0
Best Local Similarity: 99.18% Mismatches: 1
Query Match: 61.41% Indels: 2
DB: 16 Gaps: 0

US-10-023-597-24 (1-311) x US-10-292-798-1069 (1-931)
Qy      70 AsPheSerPheSerThrThrIleIleProIleMetSerPheValSerArgIys 89
Db      3 GATTTTAGTTTCTACAGACATCAATCCCAATATCTAGATGTTGTTCTTACAGAGAG 62
Qy      90 AsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhePheCysPhePheValPhe 109
Db      63 AACATTATTCTTCAACAGGCTGATAGTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 122
Qy      110 SerGlySerPheIleLeuSerAlaMetValGlu-AspArgTyrValGlyIleCysAsnPr 129
Db      123 TCTGAGCTCTTACATCTGTCGGGATGATGTA-GGACCGCTACGTGGCATCTGTAAACC 181
Qy      129 OLeuLeuYrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuGlyValTy 149
Db      182 ACTGTTGTAACAGATACCAATGTCCTCCAGGTGTTGCTCTTACTGAGGTGCTA 241
Qy      149 rGlyMetGlyValPheGlyAlaValAlaIshisThrGlyAsnIleValPheLeuThrPheCy 169
Db      242 CCGGATGGGGGCTTTTGGGGGCTGTGGCTCATACAGAAATATAGTGTTCACCTTTTG 301
Qy      169 sAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGlyLeuSerCy 189
Db      302 TGCAGACAACTTGCATCACTACATGATGACATCTCCCTCCCTTGTGAGCTCTCTG 361
Qy      189 sAsnGlySerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyVa 209
Db      362 CAACGGCTCTTACATAATGTCCTGCTCATACAGGAAATATAGTGTTCACCTTTTG 421
Qy      209 lProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSe 229
Db      422 GCCCATGTTGCGCTTTTATCTTATGCTTTTATCTTTCACACATCTCCCGCGTTAG 481
Qy      229 rSerIleGlyValArgSerIleValPheSerSerCysSerSerTyrIleIleAlaValSe 249
Db      482 TTCTGCTGAGGCGAGGTCATAAGCCTTCAGTACCTGCACTCTTCATATATTGCAATTTC 541
Qy      249 rLeuPhePheGlySerGlyValaPheThrTyrLeuIleLeuSerSerIleLeuArgValSe 269
Db      542 TCTTTTCTTTGGGTGAGGAGCTTTTACGTAACCTCAACCCCTTCATTTTACCCCTGGA 601
Qy      269 pGlnGlyIleValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuI 289
Db      602 CCAAGGGAAGTGTCTCTCCCTGTTATACCACTGAGGCGCCATGTTTAAACCATTTAT 661
Qy      289 eTyrSerLeuArgAsnIleAspValIleLeuValIlePheIleValThrPheSerArgIleSe 309
Db      662 CTACAGCCTTGAGGAAATAGGATGTCAAACTTGCCCTGAAGAGAACCTTTTCCAGATTAAG 721
Qy      309 rPheSer 311
Db      722 CTTTCT 728

RESULT 4
US-10-017-161-1299
; Sequence 1299, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS

```

FILE REFERENCE: 08435/0152  
 CURRENT APPLICATION NUMBER: US/10/017,161  
 CURRENT FILING DATE: 2002-12-18  
 PRIOR APPLICATION NUMBER: JP 2001/246789  
 PRIOR FILING DATE: 2001-06-18  
 NUMBER OF SEQ ID NOS: 2430  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 1299  
 LENGTH: 886  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: source  
 LOCATION: (1)..(886)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (201)..(686)  
 US-10-017-161-1299

Alignment Scores:  
 Pred. No.: 6,97e-30 Length: 886  
 Score: 40.00 Matches: 40  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 12.86% Indels: 0  
 Gaps: 0  
 DB: 15

US-10-023-597-24 (1-311) x US-10-017-161-1299 (1-886)

QY 28 PhePheLeuPheLeuGlyPheTyrAlaValThrValAlaGlyAsnLeuGlyLeuIleIle 47  
 DB 282 TTTCTCCGTTTCTAGGTTTCTACGCGGTCAAGGGGGGAACCTGGCTTGATATC 341  
 QY 48 LeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSer 67  
 DB 342 CTGATAGGCTCAACTCTCGCTGCATATCCCCCATGACITTTTCCCTTCAACTTGTCC 401

RESULT 5

US-10-292-798-1067  
 Sequence 1067, Application US/10292798  
 Publication No. US20030235833a1  
 GENERAL INFORMATION:  
 APPLICANT: SUMA, MAKIKO  
 APPLICANT: ASAI, KIYOSHI  
 APPLICANT: AKIYAMA, YUTAKA  
 APPLICANT: ABEURATANI, HIROYUKI  
 TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS  
 FILE REFERENCE: 08435/166  
 CURRENT APPLICATION NUMBER: US/10/292,798  
 CURRENT FILING DATE: 2002-11-13  
 PRIOR APPLICATION NUMBER: 10/017,161  
 PRIOR FILING DATE: 2001-12-18  
 PRIOR APPLICATION NUMBER: JP 2001-246789  
 PRIOR FILING DATE: 2001-06-18  
 NUMBER OF SEQ ID NOS: 2070  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 1067  
 LENGTH: 886  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: source  
 LOCATION: (1)..(886)  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (201)..(686)  
 US-10-292-798-1067

Alignment Scores:  
 Pred. No.: 6.97e-30 Length: 886  
 Score: 40.00 Matches: 40  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 12.86% Indels: 0  
 DB: 16 Gaps: 0

US-10-023-597-24 (1-311) x US-10-292-798-1067 (1-886)

QY 28 PhePheLeuPheLeuGlyPheTyrAlaValThrValAlaGlyAsnLeuGlyLeuIleIle 47  
 DB 282 TTTCTCCGTTTCTAGGTTTCTACGCGGTCAAGGGGGGAACCTGGCTTGATATC 341  
 QY 48 LeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSer 67  
 DB 342 CTGATAGGCTCAACTCTCGCTGCATATCCCCCATGACITTTTCCCTTCAACTTGTCC 401

RESULT 6

US-09-747-155-366  
 Sequence 366, Application US/09747155  
 Patent No. US20020151692A1  
 GENERAL INFORMATION:  
 APPLICANT: Rouquier, Sylvie  
 APPLICANT: Giorgi, Dominique  
 TITLE OF INVENTION: No. US20020151692A1 Polypeptides and Nucleic Acids Encoding Same  
 FILE REFERENCE: 19904-008 (C009B6834US)  
 CURRENT APPLICATION NUMBER: US/09/747,155  
 CURRENT FILING DATE: 2000-12-21  
 PRIOR APPLICATION NUMBER: 60/171,746  
 PRIOR FILING DATE: 1999-12-22  
 NUMBER OF SEQ ID NOS: 431  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO: 366  
 LENGTH: 487  
 TYPE: DNA  
 ORGANISM: Saimiri boliviensis  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (1)..(487)  
 OTHER INFORMATION: Taxon = 27679; gene = SBO222; Accession DBJ/EMBL/GenBank = AF1796  
 OTHER INFORMATION: Product = olfactory receptor  
 US-09-747-155-366

Alignment Scores:  
 Pred. No.: 3.66e-23 Length: 487  
 Score: 33.00 Matches: 33  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.61% Indels: 0  
 Gaps: 0  
 DB: 9

US-10-023-597-24 (1-311) x US-09-747-155-366 (1-487)

QY 172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGlyLeuSerCysAsnGly 191  
 DB 146 AACCTTGTCATCACTAATGATGACATCTTCCCTCTTGAGCTTCTCTGCAATGCC 205  
 QY 192 SerTyrIleAsnValLeuValIlePheIleValValThr 204  
 DB 206 TCTTACATATAATGTTCTGTCATCTTATTTGTTGTGAC 244

RESULT 7

US-09-747-155-232  
 Sequence 232, Application US/09747155  
 Patent No. US20020151692A1  
 GENERAL INFORMATION:  
 APPLICANT: Rouquier, Sylvie  
 APPLICANT: Giorgi, Dominique  
 TITLE OF INVENTION: No. US20020151692A1 Polypeptides and Nucleic Acids Encoding Same  
 FILE REFERENCE: 19904-008 (C009B6834US)  
 CURRENT APPLICATION NUMBER: US/09/747,155  
 CURRENT FILING DATE: 2000-12-21  
 PRIOR APPLICATION NUMBER: 60/171,746  
 PRIOR FILING DATE: 1999-12-22

Alignment Scores:  
 Pred. No.: 6.97e-30 Length: 886  
 Score: 40.00 Matches: 40  
 Percent Similarity: 100.00% Conservative: 0

```

; NUMBER OF SEQ ID NOS: 431
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 232
; LENGTH: 485
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(485)
; OTHER INFORMATION: Taxon = 9606; gene = HSA18; pseudogene; Accession DDBJ/EMBL/GenBank
US-09-747-155-232

Alignment Scores:
Pred. No.: 3,32e-10 Length: 485
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 9 Gaps: 0

US-10-023-597-24 (1-311) x US-09-747-155-232 (1-485)

QY 171 ABAAPAELEUVALASPHISTYMETCYASAPLLEAUPROLEUENGLULEUSERCYASN 190
DB 142 GACAACTTGTCATGATCACTAATGATGACATCTTCTCTTGAGCTCTCTGCAAC 201

RESULT 8
US-09-886-055-50
; Sequence 50, Application US/09886055
; Patent No. US20020132273A1
; GENERAL INFORMATION:
; APPLICANT: STRYER, LUBERT
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: RECEPTOR FINGERPRINTING, SENSORY PERCEPTION, AND
; FILE OF INVENTION: BIOSENSORS OF CHEMICAL SENSANTS
; FILE REFERENCE: 078003-0277150
; CURRENT APPLICATION NUMBER: US/09/886,055
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: 60/213,812
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 522
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-886-055-50

Alignment Scores:
Pred. No.: 6,36e-10 Length: 933
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 9 Gaps: 0

US-10-023-597-24 (1-311) x US-09-886-055-50 (1-933)

QY 79 PROLYSMETLEUSETSERPHEVALSERARGLYASNNILLESERPHETHRGLYCYSMET 98
DB 232 CCCAAATGCTGATGATTTTGTCACAGGAACATCATTTCTTCACAGGGTGATG 291

RESULT 9
US-09-804-291-50
; Sequence 50, Application US/09804291
; Publication No. US20030088059A1
; GENERAL INFORMATION:
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: HUMAN OLFACTORY RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: P 0278005
; CURRENT APPLICATION NUMBER: US/09/804,291
; CURRENT FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/188,914

```

```

; PRIOR FILING DATE: 2000-03-13
; PRIOR APPLICATION NUMBER: 60/192,033
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/198,474
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/199,335
; PRIOR FILING DATE: 2000-04-24
; PRIOR APPLICATION NUMBER: 60/207,702
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/213,849
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,534
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: 60/230,732
; PRIOR FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 60/266,862
; PRIOR FILING DATE: 2001-02-07
; NUMBER OF SEQ ID NOS: 529
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-804-291-50

Alignment Scores:
Pred. No.: 6,36e-10 Length: 933
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 10 Gaps: 0

US-10-023-597-24 (1-311) x US-09-804-291-50 (1-933)

QY 79 PROLYSMETLEUSETSERPHEVALSERARGLYASNNILLESERPHETHRGLYCYSMET 98
DB 232 CCCAAATGCTGATGATTTTGTCACAGGAACATCATTTCTTCACAGGGTGATG 291

RESULT 10
US-10-343-650A-395
; Sequence 395, Application US/10343650A
; Publication No. US20040067499A1
; GENERAL INFORMATION:
; APPLICANT: HACH, TATSUYA
; TITLE OF INVENTION: NOVEL G-PROTEIN COUPLED RECEPTOR
; FILE REFERENCE: 31671-186347
; CURRENT APPLICATION NUMBER: US/10/343,650A
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: JP 2000/237818
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: JP 2001/34434
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 694
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 395
; LENGTH: 933
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(933)
US-10-343-650A-395

Alignment Scores:
Pred. No.: 6,36e-10 Length: 933
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.43% Indels: 0
DB: 13 Gaps: 0

US-10-023-597-24 (1-311) x US-10-343-650A-395 (1-933)

```

## \* RESULT 11

US-10-220-382-35  
Sequence 35, Application US/10220382  
Publication No. US2003011911A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: LAU, Preeti

APPLICANT: TANG, Y. Tom

APPLICANT: PATTERSON, Chandra

APPLICANT: YAO, Monique G.

APPLICANT: SHIH, Leo D.

APPLICANT: TRIBOULET, Catherine

APPLICANT: LU, Dzung Aina M.

APPLICANT: KHAN, Farrah A.

APPLICANT: YUE, Henry

APPLICANT: POLICKY, Jennifer L.

APPLICANT: AU-YOUNG, Janice

APPLICANT: YANG, Junming

APPLICANT: HARLAND, Lee

APPLICANT: WALSH, Roderick T.

APPLICANT: LO, Terence P.

APPLICANT: BOROMSKY, Mark L.

TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS

FILE REFERENCE: PI-0044 PCT

CURRENT APPLICATION NUMBER: US/10/220,382

CURRENT FILING DATE: 2001-03-01

PRIOR APPLICATION NUMBER: 60/186,854; 60/188,384; 60/190,453; 60/190,730

PRIOR FILING DATE: 2000-03-03; 2000-03-10; 2000-03-17; 2000-03-20

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PERL Program

SEQ ID NO 35

LENGTH: 933

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US2003011911A1 7472439CBI

US-10-220-382-35

Alignment Scores:

Pred. No.: 6.36e-10

Score: 20.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 6.43%

DB: 15

Gaps: 0

US-10-023-597-24 (1-311) x US-10-220-382-35 (1-933)

Qy 79 ProlysmetLeuMetSerPheValSerArglyAsnIleIleSerPheThrGlyCysMet 98

Db 232 CCAGAAATGCTGATGAGTTTGTCTCAAGAGAAACATCTTCTCTTCAAGGGGTATG 291

RESULT 12

US-10-005-041A-13

Sequence 13, Application US/10005041A

Publication No. US2003023331A1

GENERAL INFORMATION:

APPLICANT: Casman, Stacie J.

APPLICANT: Padigaru, Muralidhara

APPLICANT: Burgess, Catherine E.

APPLICANT: Shimkets, Richard A.

APPLICANT: Spytek, Kimberly A.

APPLICANT: Gilbert, Jennifer A.

APPLICANT: Mayotte, Jane B.

APPLICANT: Baumgartner, Jason C.

APPLICANT: Mishra, Vishnu

APPLICANT: Verne, Corine AM

APPLICANT: Dickinson, Kevin S

APPLICANT: Ballinger, Robert A

APPLICANT: Wolenc, Adam R

APPLICANT: Edinger, Shlomit R

APPLICANT: MacDougall, John R

APPLICANT: Smithson, Glenda

APPLICANT: Ellerman, Karen

APPLICANT: Stone, David J

APPLICANT: Gunther, Erik

APPLICANT: Gerlach, Valerie

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: 21402-215

CURRENT APPLICATION NUMBER: US/10/005,041A

CURRENT FILING DATE: 2001-12-04

PRIOR APPLICATION NUMBER: 60/251,459

PRIOR FILING DATE: 2000-12-05

PRIOR APPLICATION NUMBER: 60/259,007

PRIOR FILING DATE: 2000-12-29

NUMBER OF SEQ ID NOS: 205

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 13

LENGTH: 953

TYPE: DNA

ORGANISM: Homo sapiens

US-10-005-041A-13

Alignment Scores:

Pred. No.: 6.5e-10

Score: 20.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 6.43%

DB: 16

Gaps: 0

US-10-023-597-24 (1-311) x US-10-005-041A-13 (1-953)

Qy 79 ProlysmetLeuMetSerPheValSerArglyAsnIleIleSerPheThrGlyCysMet 98

Db 247 CCAGAAATGCTGATGAGTTTGTCTCAAGAGAAACATCTTCTCTTCAAGGGGTATG 306

RESULT 13

US-10-024-399-1

Sequence 1, Application US/10024399

Publication No. US20030100491A1

GENERAL INFORMATION:

APPLICANT: Padigaru, Muralidhara

APPLICANT: Kekuda, Ramesh

APPLICANT: Colman, Steven D.

APPLICANT: Spytek, Kimberly A.

APPLICANT: Ballinger, Robert A.

APPLICANT: Verne, Corine A.M.

APPLICANT: Li, Li

APPLICANT: Shenoy, Suresh G.

APPLICANT: Casman, Stacie J.

TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: 21402-224AE

CURRENT APPLICATION NUMBER: US/10/024,399

CURRENT FILING DATE: 2001-12-18

PRIOR APPLICATION NUMBER: 60/256,635

PRIOR FILING DATE: 2000-12-18

PRIOR APPLICATION NUMBER: 60/259,743

PRIOR FILING DATE: 2001-01-04

PRIOR APPLICATION NUMBER: 60/299,327

PRIOR FILING DATE: 2001-06-19

PRIOR APPLICATION NUMBER: 60/261,498

PRIOR FILING DATE: 2001-01-12

PRIOR APPLICATION NUMBER: 60/263,689

PRIOR FILING DATE: 2001-01-24

PRIOR APPLICATION NUMBER: 60/267,464

PRIOR FILING DATE: 2001-02-08

PRIOR APPLICATION NUMBER: 60/271,021

PRIOR FILING DATE: 2001-02-22

PRIOR APPLICATION NUMBER: 60/275,946

PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/278,150  
PRIOR FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: 60/285,718  
PRIOR FILING DATE: 2001-04-23  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 1  
LENGTH: 976  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-024-399-1

Alignment Scores:  
Pred. No.: 6,65e-10 Length: 976  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6,43% Indels: 0  
DB: 15 Gaps: 0

US-10-023-597-24 (1-311) x US-10-024-399-1 (1-976)

Qy 79 ProlysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrgIyCysMet 98  
Db 242 CCCAAATGCTGATGAGTTTGTCTCAAGAGAACATCATTTCTTCACAGGGGTATG 301

RESULT 14  
US-10-017-161-325

Sequence 325, Application US/10017161  
GENERAL INFORMATION:  
APPLICANT: SUMA, MAKIKO  
APPLICANT: ASAI, KIYOSHI  
APPLICANT: AKIYAMA, YUTAKA  
APPLICANT: ABURATANI, HIROYUKI  
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS  
FILE REFERENCE: 084335/0152  
CURRENT APPLICATION NUMBER: US/10/017,161  
CURRENT FILING DATE: 2002-12-18  
PRIOR APPLICATION NUMBER: JP 2001/246789  
PRIOR FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 2430  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 325  
LENGTH: 1333  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: source  
LOCATION: (1)..(1333)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (201)..(1133)  
US-10-017-161-325

Alignment Scores:

Pred. No.: 9,06e-10 Length: 1333  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6,43% Indels: 0  
DB: 15 Gaps: 0

US-10-023-597-24 (1-311) x US-10-017-161-325 (1-1333)

Qy 79 ProlysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrgIyCysMet 98  
Db 432 CCCAAATGCTGATGAGTTTGTCTCAAGAGAACATCATTTCTTCACAGGGGTATG 491  
RESULT 15  
US-10-292-798-291

Sequence 291, Application US/10292798  
Publication No. US2003025833A1  
GENERAL INFORMATION:  
APPLICANT: SUMA, MAKIKO  
APPLICANT: ASAI, KIYOSHI  
APPLICANT: AKIYAMA, YUTAKA  
APPLICANT: ABURATANI, HIROYUKI  
TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS  
FILE REFERENCE: 084335/166  
CURRENT APPLICATION NUMBER: US/10/292,798  
CURRENT FILING DATE: 2002-11-13  
PRIOR APPLICATION NUMBER: 10/017,161  
PRIOR FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: JP 2001-246789  
PRIOR FILING DATE: 2001-06-18  
NUMBER OF SEQ ID NOS: 2070  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 291  
LENGTH: 1333  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
LOCATION: source  
FEATURE:  
LOCATION: (1)..(1333)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (201)..(1133)  
US-10-292-798-291

Alignment Scores:

Pred. No.: 9,06e-10 Length: 1333  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6,43% Indels: 0  
DB: 16 Gaps: 0

US-10-023-597-24 (1-311) x US-10-292-798-291 (1-1333)

Qy 79 ProlysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrgIyCysMet 98  
Db 432 CCCAAATGCTGATGAGTTTGTCTCAAGAGAACATCATTTCTTCACAGGGGTATG 491

Search completed: September 30, 2004, 17:26:19  
Job time : 550 secs



***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 08:16:15 ; Search time 4423 Seconds

(without alignments)  
3047.634 Million cell updates/sec

Title: US-10-023-597-24

Perfect score: 1577  
Sequence: 1 MAANSSSVTEFIAGLIHQ.....LRNKDVKLALKRTFSRSFS 311

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DPV=xlh  
-Q=/cgn2.1/USFTO.spool/US10023597/runat.29092004.163409.14879/app.query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastag -SUFFIX=rge -MTNMATCH=0.1 -LOOEXT=0 -LOOEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -XODE=LOCAL  
-CUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10023597@cgn.1.1.3731@runat.29092004.163409.14879 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_hlg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_ov: \*  
22: em\_or: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1577	100.0	963	6	AX555964 Sequence
2	1568	99.4	994	6	AX241500 Sequence
3	1568	99.4	172921	2	AC083958 Homo sapi
4	1568	99.4	192926	9	AP000916 Homo sapi
5	1333.5	84.6	935	6	AX241493 Sequence
6	1309	83.0	936	10	AY073678 Mus muscu
7	1309	83.0	936	10	AY318059 Mus muscu
8	1309	83.0	201304	10	AC096785 Mus muscu
9	1303	82.6	258873	2	AC096935 Rattus no
10	1273.5	80.8	930	6	AX242246 Sequence
11	1273.5	80.8	933	6	AX244609 Sequence
12	1273.5	80.8	933	6	AX448443 Sequence
13	1273.5	80.8	933	6	BD144473 Novel G-p
14	1273.5	80.8	991	6	AX702842 Sequence
15	1273.5	80.8	1333	6	AX646099 Sequence
16	1273.5	80.8	1333	9	AB065834 Homo sapi
17	1273.5	80.8	150847	9	AP001524 Homo sapi
18	1273.5	80.8	172991	2	AC083958 Homo sapi
19	1268	80.4	976	6	AX554206 Sequence
20	1266.5	80.3	953	6	AX554459 Sequence
21	1232.5	78.2	933	10	AY318057 Mus muscu
22	1230.5	78.0	933	10	AY073346 Mus muscu
23	1230.5	78.0	235933	2	AC131161 Rattus no
24	1224.5	77.6	253120	2	AC097099 Rattus no
25	1223.5	77.6	933	10	AY318056 Mus muscu
26	1223.5	77.6	209420	2	AC109178 Mus muscu
27	1221	77.4	931	6	AX646877 Sequence
28	1220.5	77.4	933	10	AF282279 Mus muscu
29	1220.5	77.4	933	10	AY073780 Mus muscu
30	1220.5	77.4	933	10	AY318060 Mus muscu
31	1220.5	77.4	209644	10	AC074314 Sequence
32	1220.5	77.4	1655	6	AX357037 Sequence
33	1217.5	77.2	235933	2	AC131161 Rattus no
34	1217.5	76.9	933	10	AY073208 Mus muscu
35	1201.5	75.4	933	10	AC114252 Rattus no
36	1189.5	75.4	933	10	AF282282 Mus muscu
37	1189.5	75.4	933	10	AY073207 Mus muscu
38	1185.5	75.2	933	10	AY318058 Mus muscu
39	1180	74.8	933	6	AX242071 Sequence
40	1180	74.8	936	6	AX244610 Sequence
41	1180	74.8	936	6	AX454190 Sequence
42	1180	74.8	936	6	AF238488 Homo sapi
43	1180	74.8	942	6	BD144474 Novel G-p
44	1180	74.8	951	6	AX448445 Sequence
45	1180	74.8	989	6	AX702946 Sequence

RESULT 1

#### ALIGNMENTS

AX555964  
 LOCUS AX555964 963 bp DNA linear PAT 27-NOV-2002  
 DEFINITION Sequence 23 from Patent WO0250275.  
 ACCESSION AX555964  
 VERSION AX555964.1 GI:25899399  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
 AUTHORS Padigaru, M., Kekuda, R., Li, L., Ballinger, R.A., Caseman, S.J., Spylek, K.A., Baumgartner, J.C. and Burgess, C.E.  
 TITLE Novel proteins and nucleic acids encoding same  
 JOURNAL Patent: WO 0250275-A 23 27-JUN-2002; Curagen Corporation (US)

FEATURES  
 source location/Qualifiers  
 1..963  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,68e-111 Length: 963  
 Score: 1577.00 Matches: 311  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

US-10-023-597-24 (1-311) x AX555964 (1-963)

QY 1 MetAlaagluAenSerSerSerValThrgupheileuAlagileuileh;sgln 20  
 Db 2 ATGGCTGCCGGAACCTCTCCCTCCGTGACAGATTATCCGAGGCTTAATCCACAG 61  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTYRAlaValThrValVal 40  
 Db 62 CCGGACCTCCAGAGTCCCGCTCTCTCTGTTCTAGGTTTCTACGGCGTCACCGCTGGT 121  
 QY 41 GlyAsnLeuGlyLeuileileuilegilyeuanSerArgLeuH;siLeProMetTYR 60  
 Db 122 GGGAACCTGGGCTTGATATCTGATAGGGCTCAACTCTCCCGTCATATCCCATGTAC 181  
 QY 61 PhePheProPheAenLeuSerLeuValAspPheSerPheSerThrTrilleleProlys 80  
 Db 182 TTTTCCCTTCAACTGTCCCTGTAGATTGTTCTCTACGACCATCATCTCCCAA 241  
 QY 81 MetLeuMetSerPheValSerArgValAsnilleleSerPheThrGlyCysMetSerGln 100  
 Db 242 ATGCTGATGAGTTTGTCTCAAGAGAAACATTTTCTTCCACAGGGTGTATGAGTCAG 301  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheileleuSerAlaMetValGlu 120  
 Db 302 TTCTTCTCTCTCTCTCTTCTTCTTCTGAGCTTCACTCGTGGGAGTGTGGAG 361  
 QY 121 AspArgTYRValGlyileCysAsnProLeuLeuTYRThrillePheMetSerProGlnVal 140  
 Db 362 GACCGCTACGTGGGCACTCTGTAACCCACTGTGTACACGATCACCAAGTCTCCACGGTG 421  
 QY 141 CysLeuLeuLeuLeuLeuGlyValTYRGLYMeGlyValPheGlyAlaValAlaH;sthr 160  
 Db 422 TGTTCCTCTCTTAACTGGGTGTCTACGGGATGGGGGTTTTTGGGGCTGTGGTCTATA 481  
 QY 161 GlyAsnileValPheLeuThrPheCysAlaAspAsnLeuValAsnH;stYrMetCysAsp 180  
 Db 482 GGAAATATATATGTTTCTCACCCTTTTGACAGACAACTGTGTCAATACATCAATGTTGAC 541  
 QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTYRilleAsnValLeuValIlePhe 200  
 Db 542 ATCTTCCCTTCTTGAAGCTCTCCTGCAAGGCTCTTACATATATGTTCTGTGATCTTT 601

QY 201 IleValValThrValGlyileGlyValProileValAlaValPheileSerTYRglyPhe 220  
 Db 602 ATGTTTGTGACCGTGGCATTTGGGGTGCCCATTTGTCGGCTTTTATCTCTTATGCTTT 661  
 QY 221 IleLeuSerSerilleleuArgValSerSerAlaGluGlyArgSerValAspPheSer 240  
 Db 662 ATTCTTCCAGCATTTCTCCCGCTTATGTTCTGTGAGGGAGGCTCTTAAAGCTTCACTAGC 721  
 QY 241 CysSerSerTYRilleleAlaValSerLeuPhePheGlySerGlyAlaPheThrTYRleu 260  
 Db 722 TGCAGCTCCACATATATGAGATTCTCTTTCTTGTGGTCAAGAGCTTTTACGTACCTC 781  
 QY 261 LysProProSerSerilleleuProLeuAspGlnGlyValSerSerLeuPheTYRThr 280  
 Db 782 AAACCCCTTCCATTTTACCCCTGACAGAGGGAAGTCTCTCCCTGTCTATACCACT 841  
 QY 281 ValValProMetPheAsnProLeuileTYRSerLeuArgAsnLysAspValIleLeuAla 300  
 Db 842 GTGGTCCCATGTTTATACCATTAATCTACAGCTGAGGAATATAGGATGTCAAATCTGCC 901  
 QY 301 LeuLysArgThrPheSerArgTlleSerPheSer 311  
 Db 902 CTGAAGAGAACCTTTTCCAGAAATTAAGCTTTTCT 934

RESULT 2  
 AX241500  
 LOCUS AX241500 994 bp DNA linear PAT 26-SEP-2001  
 DEFINITION Sequence 248 from Patent WO0127158.  
 ACCESSION AX241500  
 VERSION AX241500.1 GI:15798375  
 KEYWORDS  
 SOURCE synthetic construct  
 ORGANISM synthetic construct  
 artificial sequences.

REFERENCE  
 1 Bellenson, J., Smith, D., Lancet, D., Giusman, G., Fuchs, T. and Yanai, I.  
 TITLE Olfactory receptor sequences  
 JOURNAL Patent: WO 0127158-A 248 19-APR-2001;  
 Discentes (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)

FEATURES  
 source location/Qualifiers  
 1..994  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="(H3897 nucleotide)"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 8.4e-111 Length: 994  
 Score: 1568.00 Matches: 310  
 Percent Similarity: 99.68% Conservative: 0  
 Best Local Similarity: 99.68% Mismatches: 1  
 Query Match: 99.43% Indels: 0  
 DB: 6 Gaps: 0

US-10-023-597-24 (1-311) x AX241500 (1-994)

QY 1 MetAlaagluAenSerSerSerValThrgupheileuAlagileuileh;sgln 20  
 Db 1 ATGGCTGCCGGAACCTCTCTCCGTGACAGATTATCCCGAGGCTTAATCCACAG 60  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTYRAlaValThrValVal 40  
 Db 61 CCGGACCTCCAGAGTCCCGCTCTCTCTGTTCTAGGTTTCTACCGGGTCAAGGTTGGT 120  
 QY 41 GlyAsnLeuGlyLeuileileuilegilyeuanSerArgLeuH;siLeProMetTYR 60  
 Db 121 GGGAACCTGGGCTTGATATCTGATAGGGCTCAACTCTCGCTGCAATATCCCATGTAC 180  
 QY 61 PhePheProPheAenLeuSerLeuValAspPheSerPheSerThrTrilleleProlys 80  
 Db 181 TTTTCCCTTCAACTGTCCCTGTAGATTGTTCTCTACGACCATCATCTCCCAA 240

```

OY      61 MetLeuMetSerPheValSerArgGlySerAsnIleIleSerPheThrGlyCysMetSerGln 100
DB      241 ATGCTGATGAGTGTCTCTCAAGAGAAACATTAATTCCTTCACAGGGGTGATAGTACG 300
OY      101 PhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
DB      301 TTCCTCTCTCTCTGTTCTCTTCTCTTCTCTGAGTCCATCCCTGTCGCAATGATGATGAG 360
OY      121 AspArgGlyValGlyIleCysAsnProLeuLeuThrIleThrMetSerProGlnVal 140
DB      361 GACCCCTACGAGGAGATCTGTACCAACCTGTTGTACAGATCAGATCATCTCCCAAGTGG 420
OY      141 CysLeuLeuLeuLeuGlyValIleGlyMetGlyValPheGlyValAlaIleHisThr 160
DB      421 TGTTCCTCTCTCTTACTGGGTGTCTACGGGATGAGGGGTTTGGGGCTGCTCATCA 480
OY      161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisThrMetCysAsp 180
DB      481 GGAATATAGTGTCTCTACCTTTGTGACAACTTGTCAATCATCATCATGTGTGAC 540
OY      181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerIleAsnValIleValIlePhe 200
DB      541 ATCCCTTCCTCTCTCTGAGCTCTCCGCAACGGCTCTTACATTAATGTCCTGGTCACTTT 600
OY      201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerIleGlyPhe 220
DB      601 ATTTGTTGACCGGTGGCATTTGGGGTGGCCATTTGTCGTTTATCTCTTATGTTT 660
OY      221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerGlyAlaPheSerSer 240
DB      661 ATTCCTTCAGCATTTCTCCGCTTACTTGTCTGAGGAGAGCTCTTAACCTTCAGTACG 720
OY      241 CysSerSerIleIleIleAlaValSerLeuPhePheGlySerGlyValAlaPheThrIleu 260
DB      721 TCGAGCTCTACATTAATTCAGTTCTCTTCTTCTTGGGTGAGGAGCTTTTACGTACCTC 780
OY      261 LysProProSerIleLeuProLeuLeuSerGlyValValSerSerLeuPheThrThr 280
DB      781 AAACCCCTCTCAATTTACCTCCGACGAGGGAAGTCTCTCTCTCTCTATACCACT 840
OY      281 ValValProMetPheAsnProLeuIleIleIleSerIleuArgAsnLysAspValIleuAla 300
DB      841 GGGTGCCCATGTTTAACTTAATCACTAATCAAGCTGAGGATAGAGATGCAACTGCGC 900
OY      301 LeuLysArgThrPheSerArgIleSerPheSer 311
DB      901 CTGAAGAGAACCTTTTCAGAAATMACTTTCT 933

RESULT 3
AC083958 172991 bp DNA linear HTG 03-MAR-2001
LOCUS Homo sapiens clone RP11-18M5, WORKING DRAFT SEQUENCE, 8 unordered
DEFINITION
pieces.
AC083958
VERSION AC083958.2 GI:13184080
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 172991)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome, clone RP11-18M5
JOURNAL Unpublished
2 (bases 1 to 172991)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choedel,Y., Colangelo,M., Collins,S., Collipare,A., Cooke,P.,
Darellan,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,R., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heatford,A., Horton,L.,

```

# TITLE JOURNAL COMMENT

Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Labrecque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,  
Macdonald,P., Margis,N., McCarthy,M., McEwan,P., McKernan,K.,  
McPheters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisanu,C., Pollara,V., Raymond,C., Rebeck,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Scojanovic,N.,  
Srouguez,C., Spencer,B., Stange-Thoman,N., Stojanovic,N.,  
Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigglio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.

Direct Submission  
Submitted (08-OCT-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 2, 2001 this sequence version replaced gi:10717225.  
All repeats were identified using RepeatMasker:  
Smith A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information  
Center project name: l11149  
Center clone name: l11149

Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 169811 bases at least Q40  
Consensus quality: 171602 bases at least Q30  
Insert size: 176000; agarose-IP  
Insert size: 172291; sum-of-coverage  
Quality coverage: 7.3 in Q20 bases; agarose-IP  
Quality coverage: 7.4 in Q20 bases; sum-of-coverage

NOTES: This is a 'working draft' sequence. It currently  
\* consists of 8 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 388: contig of 388 bp in length  
1 389 488: gap of 100 bp  
\* 489 3383: contig of 2895 bp in length  
\* 3384 3483: gap of 100 bp  
\* 3484 6775: contig of 3292 bp in length  
\* 6776 6875: gap of 100 bp  
\* 6876 68914: contig of 63039 bp in length  
\* 68915 70014: gap of 100 bp  
\* 70015 87665: contig of 17651 bp in length  
\* 87666 112075: contig of 24310 bp in length  
\* 112076 112175: gap of 100 bp  
\* 112176 166387: contig of 54212 bp in length  
\* 166388 166487: gap of 100 bp  
\* 166488 172991: contig of 6504 bp in length.

## FEATURES

source

1..172991  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-18M5"  
/clone\_lib="RP11 Human Male BAC"  
1..388  
/note="assembly\_fragment  
clone\_end:SP6

misc\_feature



Db 102640 CCGGAGCTCAGGTCGCCGCTCTTCTTCTAGGTTTCTACGCGGTACCGGTG 102581

Qy 41 GYAAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db 102580 GGGAACTGGGGCTTATATCTGATAGAGGCTCAACTCTCCCTGATATCCCATGTAC 102521

Qy 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIleProLys 80

Db 102520 TTTTCCCTTCAACTTCTCCCTGATAGTTTATTTCTCTACGACCATCATTCACAAA 102461

Qy 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db 102460 ATGCTGATGAGTTTGTCTCAGAGAAACATATTTCTTCCACAGGGGTGATAGTCAG 102401

Qy 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120

Db 102400 TTCTCTTCTTCTGTTTCTTTTGTCTTTCTGAGTCTTCTCATCTGTCGGGATGAGTGTAG 102341

Qy 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db 102340 GACCCCTAGCGGCGATCTGTAACCCAGCTGTGTACAGATCACCATGTCTCCACAGTG 102281

Qy 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db 102280 TGTTCCTCTTCTTCTGAGGTCTACGCGGATGGGGTCTTTGGGGCTGTGGCTCATPCA 102221

Qy 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db 102220 GGAATATATAGTTTCTCACCTTTGTGACAGCAACCTTGTCAATCATCATATGTGTAC 102161

Qy 181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db 102160 ATCTTCCCTCTTGAAGTCTTCTGCAAGCGCTCTTACATTAATAGTCTGTGATCTTT 102101

Qy 201 IleValIleThrValGlyIleGlyValProIleValIleValPheIleSerTyrGlyPhe 220

Db 102100 ATTGTTGTACCGTGGGATGTGGGGTCCCATTTGTGCGTTTATCTTATACGTTT 102041

Qy 221 IleLeuSerSerIleLeuArgValSerSerAlaGlyGlyArgSerValAlaPheSerSer 240

Db 102040 ATCTTCTTCAGATCTCCGCGTTAGTCTGTGAGGCGAGGTCTTAAGCCTTCAGTAGC 101981

Qy 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db 101980 TGCAGCTCTCATATATTCAGATTTCTTCTTCTTGGGTCAAGGCTTTTACGTACTC 101921

Qy 261 LysProProSerIleLeuProLeuAspGlyGlyValSerSerLeuPheTyrThrThr 280

Db 101920 AAAACCCCTTCCATTTTACCCCTGACACGAGGAAAGTGTCTCTCTCTTCTATCCACT 101861

Qy 281 ValIleValProMetPheAsnProLeuIleTyrSerLeuArgAsnValAspValIleLeuAla 300

Db 101860 GGGGTGCGCATGTTTATACCATTAATCTTACAGCTCTGAGAAATAGATGTAACATTTGCC 101801

Qy 301 LeuLysArgThrPheSerArgIleSerPheSer 311

Db 101800 CTGAGAGAACCTTTTCCAGATTAAGCTTTTCT 101768

RESULT 5

AX241493

LOCUS AX241493 935 bp DNA linear PAT 26-SEP-2001

DEFINITION Sequence 241 from Patent WO0127158.

ACCESSION AX241493

VERSION AX241493.1 GI:15798368

KEYWORDS

SOURCE

ORGANISM

synthetic construct

artificial sequences.

REFERENCE

1 Bellenson, D., Smith, D., Lancel, D., Glusman, G., Fuchs, T. and Yanai, I.

TITLE Olfactory receptor sequences

JOURNAL Patent: WO 0127158-A 241 19-APR-2001;

FEATURES

source

1. 935

/organism="synthetic construct"

/mol\_type="unassigned DNA"

/db\_xref="taxon:32630"

/note="(H3890 nucleotide)"

ORIGIN

Alignment Scores:

Pred. No.: 5,79e-93 Length: 935

Score: 1333.50 Matches: 257

Percent Similarity: 92.188 Conservative: 26

Best Local Similarity: 83.718 Mismatches: 23

Query Match: 84.564 Indels: 2

DB: Gaps: 1

US-10-023-597-24 (1-311) x AX241493 (1-935)

Qy 1 MetAlaIleGluAsnSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20

Db 1 ATGCTGCAAGAACTCTCTCTCGTACAGCGTTATCTCTCAGGCTTACCGACAG 60

Qy 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrVal 40

Db 61 CCGGAGCTCCAGATCCCGCGCTTCTCTGTTCTTACGTTTCTACCGCGGTACAGTGTG 120

Qy 41 GlyAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

Db 121 GGAACCTGGGCTGTGATATCCGATAGGCTTACCTCGCTGATATATCCCAATGATC 180

Qy 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIleProLys 80

Db 181 TTTTCCCTTCACTTGTCTTCTCATAGATTTAGTATATCCATCCCTCGCCCTTAA 240

Qy 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100

Db 241 ATGCTGATGAGCTTTGTGCTCA---GAGAACATCATTTCCATACAGGATGATGACTCAG 297

Qy 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120

Db 298 CTTTCTTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 357

Qy 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

Db 358 GACCGCTACGTGGGACATCTGTAACCATGTTTACACGCTTACACGCTTACGCTTCCAGATG 417

Qy 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

Db 418 TGTTCCTCTTCTTCTGAGGTGTCTATGGGATGGGAT-TTTGGGCTGTGGCTCATATG 476

Qy 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

Db 477 GGAACATATAGTTTATGCTCTTTGTGAGAGAACCTTCTCATATATGCTGTGAC 536

Qy 181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

Db 537 ATCTTCTCTCTTGAAGCTCTCTGCAACAGCTTACATTAATTTGCTGTGGGTGTTT 586

Qy 201 IleValIleThrValGlyIleGlyValProIleValIleValPheIleSerTyrGlyPhe 220

Db 597 ATATATGACCGTGTGATGGGTGCGCATGTCCACATTTTCTCTTATGGTTT 656

Qy 221 IleLeuSerSerIleLeuArgValSerSerAlaGlyGlyArgSerValAlaPheSerSer 240

Db 657 ATCTTCTTCCAGATCTTCCATTAATGTTCCACAGAGGAGGTCTTAAGCCTTCAGATAC 716

Qy 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db 717 TGCAGTTCACATTAATGTTGATGCTTTCTTGGGTCAAGGCTTTCATGATGATCTC 776

Qy 261 LysProProSerIleLeuProLeuAspGlyGlyValSerSerLeuPheTyrThrThr 280

Db 777 AACCACTTATTCTTACCCCTGACGAGGAGGAAAAGTCTTCATTTTTGTACTGCT 836  
 Qy 281 ValValPrometPheasnProleuileTySerleuArgAsnLysAspValylsLeuAla 300  
 Db 837 GTGGTGGCCCATGTTTAAACCATTAATCTACAGCCTGAGAGATTAAGATGTAAGATTGCC 896  
 Qy 301 LeuysArgThrPheSerArg 307  
 Db 897 CTGAGAGAACTTTTGCA 917  
 RESULT 6  
 AY073678 936 bp DNA linear ROD 03-APR-2002  
 LOCUS Mus musculus olfactory receptor MOR161-5 gene, complete cds.  
 DEFINITION AY073678  
 ACCESSION AY073678.1 GI:18480653  
 VERSION  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 936)  
 AUTHORS Zhang, X. and Firestein, S.  
 TITLE The olfactory receptor gene superfamily of the mouse  
 JOURNAL Nat. Neurosci. 5 (2), 124-133 (2002)  
 MEDLINE 21676863  
 PUBMED 11802173  
 REFERENCE 2 (bases 1 to 936)  
 AUTHORS Young, J.M., Friedman, C., Williams, E.M., Ross, J.A., Tonnes-Priddy, L. and Trask, B.J.  
 TITLE Different evolutionary processes shaped the mouse and human olfactory receptor gene families  
 JOURNAL Hum. Mol. Genet. 11 (5), 535-546 (2002)  
 MEDLINE 21864068  
 PUBMED 11875048  
 REFERENCE 3 (bases 1 to 936)  
 AUTHORS Adams, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
 FEATURES  
 source location/Qualifiers  
 1..936  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="9"  
 /map="9A5"  
 /note="GA\_X5UBB7W60AJ-1756545-1755610"  
 <1..>936  
 /product="Olfactory receptor MOR161-5"  
 1..936  
 /codon\_start=1  
 /evidence=not experimental  
 /product="Olfactory receptor MOR161-5"  
 /protein\_id="AA161341.1"  
 /db\_xref="GI:18480654"  
 /translation="MAAANTSSVAEFLVLVGLDQDQLPFLFLFGFYIVTMVGNLGLTILGNSHLHPYFELFNLSIFIDFSYTLTPKLVGFVLRKNIISYAGMDTFPFCFVFSBSYLIISAMAYDRYVAICFKPLSYTMSPOVCSYLSGVYGMVGAVAHMGNLOPISFCADNIINRYMCDIIPLEISCSNYSINLIVAVITIGIVITIFISYGFILSLIHSISKEGSKAPSTCTSHIIVYLFPGSGAFMYLKPSPSLDQKXVSSV FYAAVVMFPLIYLSLNKDVAKIALKTKLSRKIFS"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.26e-91 Length: 936  
 Score: 1309.00 Matches: 246  
 Percent Similarity: 89.07% Conservative: 31  
 Best Local Similarity: 79.10% Mismatches: 34  
 Query Match: 83.01% Indels: 0  
 Ds: 10 Gaps: 0  
 US-10-023-597-24 (1-311) x AY073678 (1-936)

Qy 1 MetAlaIaGlUaSnSerSerSerValThnGiUpheileLeuAlaGlyLeuileHisGln 20  
 Db 1 ATGGCTGACGCGAACAACCTCTCAATGAGGAGAGTTCATCTTGTGGATTACACACGAG 60  
 Qy 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheThyAlaValThyValVal 40  
 Db 61 CCTCAACTCCAAATCCCACTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 120  
 Qy 41 GlyAsnLeuGlyLeuile 60  
 Db 121 GGGAACTGGGCTTATCATCATGATAGGCGTGAAGCTTCACCTTCATATCCCATGATC 180  
 Qy 61 PhePheProPheAsnLeuSerleuValAspPheSerPheSerThrThrileleProlys 80  
 Db 181 TTTTTCCTTCCTCAATTTGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 240  
 Qy 81 MetLeuMetSerPheValSerArgLysAsnlelelelelelelelelelelelelelelelele 100  
 Db 241 ATGCTGGTGGGCTTTGTTCTTGAGAGAAACATCATCTCTATGCGAGGCTGATGACACAG 300  
 Qy 101 PhePhePhePheCysPhePheValPheSerGlySerPheilelelelelelelelelelelele 120  
 Db 301 TTTTCTTCTCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360  
 Qy 121 AspArgTrpValGlyileCysAsnProleuLeuThrThrilelelelelelelelelelelele 140  
 Db 361 GACCGCTATGTCACATCTGTAAACCACTGTGATAGTGCACCACTGCTCCCAATG 420  
 Qy 141 CysLeuLeuLeuLeuGlyValTyrglymeGlyValPheGlyAlaValAlaHisThr 160  
 Db 421 TTTTCAATCTTTGTCAGGTGCTATGATGAGAGTGGAGTGGGCTGGGCTGGGCTGGGCTGGG 480  
 Qy 161 GlyAsnle 180  
 Db 481 GGAAACCTACATTTATAGCTTCTGCTGTACACACATTAACCACTATATATGTTGTAC 540  
 Qy 181 IleleuProleuLeuGlyLeuSerCysAsnGlySerTyrllelelelelelelelelelelele 200  
 Db 541 ATATATCCCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 Qy 201 IleValValThrValGlyileGlyValProlelelelelelelelelelelelelelelelele 220  
 Db 601 ATGTTGAGACCATGAGCATGAGTGCATGCTTCAATCAATCAATCAATCAATCAATCAATCAAT 660  
 Qy 221 IleleuSerSerIleleuArgValSerSerIlelelelelelelelelelelelelelelelele 240  
 Db 661 ATCTCTCTGAGCTTCTTCACATTAATGTTTAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720  
 Qy 241 CysSerSerTyrlle 260  
 Db 721 TGTACGTCCCATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780  
 Qy 261 LysProProSerIleleuProleuAspGlnGlyValSerSerlelelelelelelelelelelele 280  
 Db 781 AAACCACTTCAATGTTGCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840  
 Qy 281 ValValPrometPheasnProleuileTySerleuArgAsnLysAspValylsLeuAla 300  
 Db 841 GTGGTGGCCCATGTTTAAACCATTAATCTACAGCCTGAGAGATTAAGATGTAAGATTGCC 900  
 Qy 301 LeuysArgThrPheSerArg 311  
 Db 901 CTGAAAGAAACCTTGACGAGAGAAAATTTCTCT 933  
 RESULT 7  
 AY318059 936 bp DNA linear ROD 14-OCT-2003  
 LOCUS Mus musculus olfactory receptor GA\_X6K0217PVD-31540342-31541277  
 DEFINITION (GA\_X6K0217PVD-31540342-31541277) gene, complete cds.  
 ACCESSION AY318059  
 VERSION AY318059.1 GI:32063397  
 KEYWORDS







Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorenz, L., Louie, H., Lozano, R., Lu, X., Ma, J.,  
 Maheshwari, M., Mahindrasekaran, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,  
 Mawlawi, S., McLeod, M., McNeill, T., Meenen, E.,  
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Muniz, N., Murphy, M., Nair, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwankwelu, O., Okunnu, G., Olarunso, A., Pal, S., Parks, K.,  
 Passerak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plopper, F., Poindexter, A., Popovic, D., Primm, E., Pu, L.,  
 Puro, M., Quiroz, J., Rachin, E., Reeves, K., Register, M., Reigh, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,  
 Sanders, W., Saverly, G., Scherer, S., Scott, G., Slatman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C., Smaj, D.,  
 Sneed, A., Sodergren, E., Song, X., Z., Sorrelle, R., Sosa, J.,  
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K.,  
 Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,  
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Wilson, R., Wleczek, R., Woodard, H., Worley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, U., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von  
 Niederhausern, A., Weiss, R., Smith, D., Holt, R., Smith, H., O.,  
 Weinstein, G., and Gibbs, R. A.  
 Direct Submission  
 2 (bases 1 to 258873)  
 Unpublished  
 Title  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Rat Genome Sequencing Consortium.  
 Direct Submission  
 Submitted (10-MAY-2003) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On May 10, 2003 this sequence version replaced gi:24953430.  
 The sequence in this assembly is a combination of BAC based reads  
 and whole genome shotgun sequencing reads assembled using Atlas  
 (http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described  
 in the feature table below represents a scaffold in the Atlas  
 assembly (a 'contig-scaffold'). Within each contig-scaffold,  
 individual sequence contigs are ordered and oriented, and separated  
 by sized gaps filled with Ns to the estimated size. The sequence  
 may extend beyond the ends of the clone and there may be sequence  
 contigs within a contig-scaffold that consist entirely of whole  
 genome shotgun sequence reads. Both end sequences and whole genome  
 shotgun sequence only contigs will be indicated in the feature  
 table.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: http://www.hgsc.bcm.tmc.edu/  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: GHNV  
 Center clone name: CH230-154L7  
 ----- Summary Statistics  
 Assembly program: Atlas 3.0  
 Consensus quality: 207030 bases at least Q40  
 Consensus quality: 209712 bases at least Q30  
 Consensus quality: 211447 bases at least Q20  
 Estimated insert size: 223149; sum-of-contigs estimation  
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces

is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1 245424: contig of 245424 bp in length  
 \* 245425 245524: gap of unknown length  
 \* 245525 246877: contig of 1353 bp in length  
 \* 246878 246977: gap of unknown length  
 \* 246978 254737: contig of 7760 bp in length  
 \* 254738 254837: gap of unknown length  
 \* 254838 258873: contig of 4036 bp in length.  
 \* Location/Qualifiers  
 1. 258873  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"  
 /clone="CH230-154L7"  
 501..1322  
 /note="clone boundary  
 clone end:17  
 site:ECORI  
 end\_sequence=B2107537"  
 243527..245424  
 /note="wgs\_contig"  
 -----  
 ORIGIN  
 misc\_feature  
 misc\_feature  
 -----  
 Alignment Scores:  
 Pred. No.: 3,972-88 Length: 258873  
 Score: 1303.00 Matches: 245  
 Percent Similarity: 88.10% Conservative: 29  
 Best Local Similarity: 78.78% Mismatches: 37  
 Query Match: 82.63% Indels: 0  
 DB: 2 Gaps: 0  
 -----  
 US-10-023-597-24 (1-311) x AC096935 (1-258873)  
 QY 1 MetAlaAlaGluAuaSerSerSerValThGluPheIleLeuAlaGlyLeuIleHicIn 20  
 Db 155892 ATGACTCGAGAAACACCTCTCGGTGACAGAGTTCATCTTGTGATTAACAGACAG 155833  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyraIaValThValaI 40  
 Db 155832 CCGACCTCCAGATCCACCTCTTATCTCTTCTCGGTTTCAAGTCACCTCTG 155773  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 Db 155772 GCGAACCTGGGCTGATCACACTGATGAGGTGAACCTCACCTGCATATTCATGATAC 155713  
 QY 61 PhePheProPheAuaIuaSerSerLeuValAspPheSerPheSerThrIleIleProLys 80  
 Db 155712 TTTTCTCTCTCACTGCTGCTCTTATGATTTCTGATGATCTTACTACACTTACCCCTTAA 155653  
 QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100  
 Db 155652 ATGCTGGTGGGCTTTGTTGGAGAGAACATATCTTCCATACGAGGCTGATGATCAG 155593  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerIleIleLeuSerAlaMetValGlu 120  
 Db 155592 TTTCTCTCTCTCTTTCTTTCTTTGTTGTTGAGTCTGACCTTGTGGCAATGGCATAC 155533  
 QY 121 AspArgTyraValGlyIleCysAsnProLeuLeuTyThrIleThrMetSerProGlnVal 140  
 Db 155532 GACCGCTATGTTGCTCATCTGTAACCATTTGTTGATAGGTCACCACTGCTCCCAAGTG 155473  
 QY 141 CysLeuLeuLeuLeuLeuGlyValTyrglyMetGlyValPheGlyAlaValAlaHisThr 160  
 Db 155472 TGCTCATCTCTTTTGTGATGAGGCTGATGAGGTGGGTTTGGGGCTGTGGCCCATATG 155413  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyMetCysAsp 180  
 Db 155412 GGAACCTTACAGTTTATTAACCTTCTGTGCTGACCAACATCATTAACCATATATGTGTGAC 155353

```

QY      181  I I l e u P r o l e u G l u l e u S e r C y s a n g l y S e r T y r I l e a s n V a l l e u V a l l e P h e 200
      155352 A T C A T T C C C T C T G A G C T G C C G A A T A G C T T T A C A T C A A T T G T G T G T T T 155293
QY      201  I l e a l V a l T h r V a l G l y l e g l y V a l P r o l l e V a l A l a V a l P h e l l e S e r T y r G l y P h e 220
      155292 A T T G T T G T A C A C A T T G G C A T T G A G T A C C A T T G T T A C C A T T T T A T T C T T A G G C T T C 155233
Db      155292 A T T G T T G T A C A C A T T G G C A T T G A G T A C C A T T G T T A C C A T T T T A T T C T T A G G C T T C
QY      221  I l e u S e r S e r I l e u e u r g V a l S e r S e r A g l u g l y A r g S e r L y s V a l P h e S e r S e r 240
      155232 A T C C T C T C T A G C A T T C C A C A T T G T T C C A A G A G G C A G G T C C A A G C C T T A G A C C 155173
Db      155232 A T C C T C T C T A G C A T T C C A C A T T G T T C C A A G A G G C A G G T C C A A G C C T T A G A C C
QY      241  C y s S e r S e r T y r I l e a l A l a V a l S e r l e u P h e p h e g l y S e r G l y A l a P h e T h r T y l e u 260
      155112 T G C A C T T C T C A T A T T A T T G T A G T G T C T C T C T T T T T G G G T C A G A G C T T T A T G T A T C T C 155113
Db      155112 T G C A C T T C T C A T A T T A T T G T A G T G T C T C T C T T T T T G G G T C A G A G C T T T A T G T A T C T C
QY      261  L y s P r o P r o S e r I l e u P r o l e u a s p g l n g l y l y s V a l S e r S e r l e u P h e T y r T h r 280
      155112 A A A C A C C T T C T A G T T G C C T C T G A T C A G G A A A A G T C C T C T G T T T C T A T A C T G C T 155053
Db      155112 A A A C A C C T T C T A G T T G C C T C T G A T C A G G A A A A G T C C T C T G T T T C T A T A C T G C T
QY      281  V a l V a l P r o m e t P h e a s n P r o l e u l e t y r S e r l e u a r g a s n l y a s p V a l l y l e u A l a 300
      155052 G T A G T G C C A T G T T A A T C A T T A T A T C T A T A G C C T G A G A A T A A G A T G T C A A A A T T G C T 154993
Db      155052 G T A G T G C C A T G T T A A T C A T T A T A T C T A T A G C C T G A G A A T A A G A T G T C A A A A T T G C T
QY      301  L e u l y s A r g T h r P h e S e r A r g l l e s e r P h e S e r 311
      154992 C T G A A G A A A A C C T T G A C A A A A A A T T T C T C T 154960
Db      154992 C T G A A G A A A A C C T T G A C A A A A A A T T T C T C T 154960

RESULT 10
AX242246
LOCUS      AX242246          930 bp      DNA          linear      PAT 26-SEP-2001
DEFINITION Sequence 994 from Patent WO0127158.
ACCESSION  AX242246
VERSION     AX242246.1 GI:15799121
KEYWORDS
SOURCE      synthetic construct
            synthetic construct
            artificial sequences.
REFERENCE   1  Bellenson, J., Smith, D., Lancelot, D., Glusman, G., Fuchs, T. and
            Yanai, I.
            TITLE      Olfactory receptor sequences
            JOURNAL    Patent: WO 0127158-A 994.19-APR-2001;
            DISCLOSURE (US); YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)
FEATURES
            source
            1..930
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="(H389844 nucleotide)"
ORIGIN
Alignment Scores:
Pred. No.:      2,14e-88      Length:      930
Score:          1273.50      Matches:      250
Percent Similarity: 87.46%      Conservative: 22
Best Local Similarity: 80.39%      Mismatches: 38
Query Match:    80.75%      Indels:      1
DB:             6      Gaps:      1

US-10-023-597-24 (1-311) x AX242246 (1-930)
QY      1  M e t A l a I a g l u a n s e r S e r V a l T h r G l u P h e l l e u a l a g l y l e u l l e H s g l n 20
      1 A T G G A G A G C A A A A A C - - - T C T T C T G T G C A G A G T T A T A C C T G A A G C C T T A A C C C A C A G 57
Db      1 A T G G A G A G C A A A A A C - - - T C T T C T G T G C A G A G T T A T A C C T G A A G C C T T A A C C C A C A G 57
QY      21  P r o g l y e u G l n V a l P r o V a l P h e P h e u P h e u G l y P h e T y r A l a V a l T h r V a l 40
      58 C C G G A C T G C G A T C C C C C T C T C T C T G G G T T C T A C A G G T C A C C G T G 117
Db      58 C C G G A C T G C G A T C C C C C T C T C T C T G G G T T C T A C A G G T C A C C G T G 117
QY      41  G l y s n l e u G l y l e u l l e l l e u l l e g l y l e u a n s e r a r g l e u H s l e P r o m e t T y r 60
      41 G l y s n l e u G l y l e u l l e l l e u l l e g l y l e u a n s e r a r g l e u H s l e P r o m e t T y r 60

```

```

Db      118  G G A A C C T G G G C T T G A T A C C C T G A T T G G C T G A A C T C C A C C T G C A C A C T C C A T G T A C 177
QY      61  P h e P h e P r o P h e a s n l e u S e r l e u V a l a p P h e S e r P h e S e r T h r l e l l e P r o l y s 80
      178 T T C T C C T T T T A A C C T C T T T A A T A G A T T C T C T T C C A C A C A C A C A C A C A C T C C C A A 237
Db      178 T T C T C C T T T T A A C C T C T T T A A T A G A T T C T C T T C C A C A C A C A C A C A C A C T C C C A A 237
QY      81  M e t l e u S e r P h e V a l S e r a r g l y a s n l l e l l e S e r P h e T h r G l y C y s e r S e r g l n 100
      238 A T G C T A T A G A G T T T G T C T C A G A A G A A C A T C A T T C C T T A C A G G G T G T A T G A C T C A G 297
Db      238 A T G C T A T A G A G T T T G T C T C A G A A G A A C A T C A T T C C T T A C A G G G T G T A T G A C T C A G 297
QY      101  P h e P h e P h e C y s P h e P h e V a l P h e S e r g l u s e r P h e l l e u S e r A l a M e t V a l g l u 120
      298 C T C T C T C T C G C T C T T T G T G C G T C C T G A G C C T T C A T C C T G C A G C A T G C G A T G C G C T A T 357
Db      298 C T C T C T C T C G C T C T T T G T G C G T C C T G A G C C T T C A T C C T G C A G C A T G C G A T G C G C T A T 357
QY      121  A s p a r g T y r V a l G l y l e C y s a n P r o l e u l y t y r l l e T h r M e t S e r P r o l n a l 140
      358 G A C C C T A C G T G C C A C T C T G A C C A C T G T T A C A C A G T A C C A G T C T T G C C A G G T 417
Db      358 G A C C C T A C G T G C C A C T C T G A C C A C T G T T A C A C A G T A C C A G T C T T G C C A G G T 417
QY      141  C y s l e u l e u l e u l e u G l y V a l T y r g l y M e t G l y V a l P h e G l y A l a V a l A h i s T h r 160
      418 T G T T G C T C T T T T G T T G G G T G C T A T A G G A T G G G T T G T G C G G C C A T G C C A C A 477
Db      418 T G T T G C T C T T T T G T T G G G T G C T A T A G G A T G G G T T G T G C G G C C A T G C C A C A 477
QY      161  G l y a s n l l e V a l P h e l e u T h r P h e C y s A l a a s p a n l e u V a l a s n l s t y r M e t C y s a s p 180
      478 G G A A G A T A T A G A A C C T G A C C T T C T G T G C A C A C C T T G C A A T C A T T C A T T C A T G T G T A C 537
Db      478 G G A A G A T A T A G A A C C T G A C C T T C T G T G C A C A C C T T G C A A T C A T T C A T T C A T G T G T A C 537
QY      181  I l e u P r o l e u G l u l e u S e r C y s a n g l y S e r T y r I l e a s n V a l l e u V a l l e P h e 200
      538 A T C C T C C T C T C T T G A G C T C C C T G C A A G C C T T T A C A T A A A G A C T G T G T C T T 597
Db      538 A T C C T C C T C T C T T G A G C T C C C T G C A A G C C T T T A C A T A A A G A C T G T G T C T T 597
QY      201  I l e a l V a l T h r V a l G l y l e g l y V a l P r o l l e V a l A l a V a l P h e l l e S e r T y r G l y P h e 220
      598 A T T G T G T G C C G T T G A C C T T G A A T G C C A T T G C A C T G C T T A A T T C T T A T G C C C T C 657
Db      598 A T T G T G T G C C G T T G A C C T T G A A T G C C A T T G C A C T G C T T A A T T C T T A T G C C C T C 657
QY      221  I l e u S e r S e r I l e u e u r g V a l S e r S e r A g l u g l y A r g S e r L y s V a l P h e S e r S e r 240
      658 A T C C T C C A G A T T C T A C A C A C A G C T T C T A C A G A G C A G C T C C A A G C C T T T A G T A C T 717
Db      658 A T C C T C C A G A T T C T A C A C A C A G C T T C T A C A G A G C A G C T C C A A G C C T T T A G T A C T 717
QY      241  C y s S e r S e r T y r I l e a l A l a V a l S e r l e u P h e p h e g l y S e r G l y A l a P h e T h r T y l e u 260
      718 T C A G T T C C A C A T A T T A T T A G T T C T C T T C T T T G T G G T C G G C C T T A T A T A T G A T C T C 777
Db      718 T C A G T T C C A C A T A T T A T T A G T T C T C T T C T T T G T G G T C G G C C T T A T A T A T G A T C T C 777
QY      261  L y s P r o P r o S e r I l e u P r o l e u a s p g l n g l y l y s V a l S e r S e r l e u P h e T y r T h r 280
      778 A A A C C C T T T C A T C C T G C C C T C A G A G G A A G T G C C T C C T C T T C T A T A C C A T A 837
Db      778 A A A C C C T T T C A T C C T G C C C T C A G A G G A A G T G C C T C C T C T T C T A T A C C A T A 837
QY      281  V a l V a l P r o m e t P h e a s n P r o l e u l e t y r S e r l e u a r g a s n l y a s p V a l l y l e u A l a 300
      838 A T A G T C C C G G T T A A C C A T T A T C T A T A T G C T T G A G A A C A A G A T G C A A A G T T G C C 897
Db      838 A T A G T C C C G G T T A A C C A T T A T C T A T A T G C T T G A G A A C A A G A T G C A A A G T T G C C 897
QY      301  L e u l y s A r g T h r P h e S e r A r g l l e s e r P h e S e r 311
      898 C T G A G A G A A C T T T G G C A G A A A A A T C T T T C T 930
Db      898 C T G A G A G A A C T T T G G C A G A A A A A T C T T T C T 930

RESULT 11
AX244609
LOCUS      AX244609          933 bp      DNA          linear      PAT 28-SEP-2001
DEFINITION Sequence 35 from Patent WO0166742.
ACCESSION  AX244609
VERSION     AX244609.1 GI:15859510
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1  Lai, P., Tang, Y.T., Patterson, C., Yao, M.G., Shih, L.L.,
            Tribouley, C.M., Lu, D.A., Yue, H., Khan, F.A., Policky, J.L.,
            Au-Yang, Y., Yang, Y., Harland, L., Walsh, R.T., Lo, T.P. and
            Borowsky, M.L.
            TITLE      G-protein coupled receptors
            JOURNAL    Patent: WO 0166742-A 35.13-SEP-2001;
            Incyte Genomics, Inc. (US)

```

FEATURES

Location/Qualifiers

1..933

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

/note="Incylce ID No: 7472439CB1"

ORIGIN

Alignment Scores:

Pred. No.: 2.15e-88 Length: 933

Score: 1273.50 Matches: 250

Percent Similarity: 87.46% Conservative: 22

Best Local Similarity: 80.39% Mismatches: 38

Query Match: 80.75% Indels: 1

DB: 6 Gaps: 1

US-10-023-597-24 (1-311) x AX244609 (1-933)

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheLeuAlaGlyLeuIleHisGln 20

1 ATGGCAGCCAAAAC--TCTTCTGTGACAGATTATCTGAGAGCTTAACCCACAG 57

QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValAl 40

58 CCGGGAGCTCGGATCCCTCTTCTCTGTTCTGGGTTTCTACACGGTCCACCGTGTG 117

QY 41 GlyAsnLeuGlyLeuIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

118 GGGACCTGGGCTTATACCTGATGGGCTGAACTCTACCTGCACATCCCATGATAC 177

QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheThrThrIleIleProlys 80

178 TCTTCTCTTTTAACCTCTCTTAATATGATTTCTGTTCTCCACATCACTCACTCCCAA 237

QY 81 MetLeuMetSerPheValSerArglyAsnIleIleSerPheThrGlyCysMetSerGln 100

238 ATGCTGATAGTTTGTCTCAAGAAACATCATTTCTTCAAGAGGTATGACTGAG 297

QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

298 CTTCTTCTTCTGCTCTTTGTGCTGCTGAGTCTTATCTCTGACGCGATGGCGAT 357

QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

358 GACCGCTACGTGGCCATCTGTAAACCATGTTGTACACAGTACCATGATCTGACGATG 417

QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

418 TGTTCCTCTCTTTGTGGGTGCTTATGGATGGGTTGCTGGGCGCATGGCCACACAC 477

QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180

478 GGAAGCATATGACCTGACCTTGTGTGCTGACCACTTGCATATTCATGATGTGAC 537

QY 181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200

538 ATCTTCTCTCTCTTGTAGCTCTCTGCAACACTCTTATGATGAAGCTGTGTCTTT 597

QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220

598 ATGTGCTGGCTGTAGCTGTGAATGCCCATGTGACATGCTTTATTTCTATGCCCTC 657

QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyValSerSerLeuAlaPheSerSer 240

658 ATCTCTCCAGCATTTCTTACCAACAGTTTCTACAGAGGAGGATCCAAAGCCCTTGTGATCT 717

QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

718 TGGAGTTCACACATATGATGTTCTTTCTTTGTTGTTCTGAGTCTTCAATGATCTC 777

QY 261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280

778 AAACCCCTTTCATCTGCGCCCTCGAGCAAGGAAAGTGTCTCCCTGTTCTATACCATTA 837

QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

838 ATATCCCTCCCGGTGTTAAACCATTAATCTATGAGCTTGAGAAACAGATGTCAAAGTTGCC 897

QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311

898 CTGAGGAAGACTTTGGGACAGAAAATCTTTCT 930

DB

RESULT 12

AX448443 933 bp DNA linear PAT 03-JUN-2002

LOCUS

DEFINITION

Sequence 109 from Patent WO0224726.

AX448443

ACCESSION

AX448443.1 GI:21697342

KEYWORDS

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1

Veithen, A.

Olfactory and Pheromones g-protein coupled receptors

Patent: WO 0224726-A 109 28-MAR-2002;

ChemCom S.A. (BE)

FEATURES

Location/Qualifiers

1..933

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

1..933

/note="unnamed protein product"

/codon\_start=1

/protein\_id="CAD37543.1"

/db\_xref="GI:21697343"

/db\_xref="REMBL:CAD37543"

/translation="MAAKNSVTEFLEIGTTHQPGLRIPFLFGLFTVTVVGNLGLITLIGNSHLPMTFPLFNLSDIDCFSTTTPKMLMSVSRKNIISFGCMTOLEFCFVVSSEFISLMAAYDRVAICNPLFSTVMSGCVCLLILGAGMGFAGMAHATG SIMNLVFCADNLVNHPCDILPLELSECNSSVNEELVVFVAVDVGMPVIVTISVALIISLISHNSTEGRSGAFSTGSHIIVLSIFGSGAFMYLKRSLILPLEGKVSLSF YTIIVPVLNPLIYSLRNKVKVALRRTLRKITS"

ORIGIN

Alignment Scores:

Pred. No.: 2.15e-88 Length: 933

Score: 1273.50 Matches: 250

Percent Similarity: 87.46% Conservative: 22

Best Local Similarity: 80.39% Mismatches: 38

Query Match: 80.75% Indels: 1

DB: 6 Gaps: 1

US-10-023-597-24 (1-311) x AX448443 (1-933)

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheLeuAlaGlyLeuIleHisGln 20

1 ATGGCAGCCAAAAC--TCTTCTGTACAGATTATCTGAGAGCTTAACCCACAG 57

QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValAl 40

58 CCGGAGCTGGGATCCCTCTTCTCTGTTCTGGGTTTCTACACGGTCCACCGTGTG 117

QY 41 GlyAsnLeuGlyLeuIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60

118 GGGACCTGGGCTTATACCTGATGGGCTGAACTCTACCTGCACATCCCATGATAC 177

QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheThrThrIleIleProlys 80

178 TCTTCTCTTTTAACCTCTCTTAATATGATTTCTGTTCTCCACATCACTCACTCCCAA 237

QY 81 MetLeuMetSerPheValSerArglyAsnIleIleSerPheThrGlyCysMetSerGln 100

238 ATGCTGATAGTTTGTCTCAAGAAACATCATTTCTTCAAGAGGTATGACTGAG 297





```

QY      261 LysProSerIleuProLeuaspGlnGlyValSerSerIleuPheThrThr 280
      |||
      778 AAACCCCTTCATCTGCCCCCGAGCAAGGAAAGTCTCCCTGTTATACCAATA 837
      |||
QY      281 ValValProMetPheAsnProLeuIleTySerIleuArgAsnLysAspValLysLeuAla 300
      |||
      838 ATAGTCCCCGTGTAAACCATTAATCTATAGCTTGAGGACAGAGATGTCAAGTTGCC 897
      |||
QY      301 LeuLysArgThrPheSerArgIleSerPheSer 311
      |||
      898 CTGAGGAGAACTTTGGCGAGAAAATCTTTCT 930
      |||
RESULT 14
AX702842      AX702842      991 bp      DNA      linear      PAT 03-APR-2003
LOCUS      Sequence 71 from Patent WO02059313.
ACCESSION      AX702842
VERSION      AX702842.1
KEYWORDS      GI:29537953
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1.
AUTHORS      Li, L., Ballinger, R.A., Padigar, M., Kekuda, R., Colman, S.D.,
      Spivek, K.A., Casman, S.J., Vernet, C.A., Shenoy, S.G., Guise, V.,
      Malpanker, U.M., Baringer, S., Gerlach, V., Smithson, G., Stone, D.J.,
      Sciore, P., MacDougall, J.R., Gunther, E., Peyman, J.A., Ellerman, K.,
      Gangoli, E.A. and Millet, I.
      G-protein coupled receptors and nucleic acids encoding same
      Patent: WO 02059313-A 71 01-AUG-2002;
      Curagen Corporation (US)
FEATURES
Source      1..991
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.:      2,299-88      Length:      991
Score:      1273.50      Matches:      250
Percent Similarity:      87.46%      Conservative:      22
Best Local Similarity:      80.39%      Mismatches:      38
Query Match:      80.75%      Indels:      1
DB:      6      Gaps:      1
US-10-023-597-24 (1-311) x AX702842 (1-991)
QY      1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
      |||
      49 ATGGCAGCCAAAAC---TCTCTGTGACAGAGTTATCTCGAAGGCTTAACCCACAG 105
      |||
QY      21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheThrAlaValThrValAl 40
      |||
      106 CCGGAGCTCGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 165
      |||
QY      41 GlyAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
      |||
      166 GGGAACTCTGGGTGATTAACCTGATGGCTGGAAGTCTCACTGCACACTCCCAATGAC 225
      |||
QY      61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleProLys 80
      |||
      226 TCTCTCTTTTAACTCTCTTTAATAGATTCTGTCTCTCCATCACTCACTCCCAAA 285
      |||
QY      81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
      |||
      286 ATGCTATATAGTTTGTCTCAAGAGAAACATCATTTCTTCAAGGGGTATAGACTGAG 345
      |||
QY      101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
      |||
      346 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 405

```

```

QY      121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
      |||
      406 GACCGTACAGGCGCATCTTAACCCACTGTGACACACTCAACATGCTTGGCCAGATG 465
      |||
QY      141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValAlaPheGlyAlaValAlaHisThr 160
      |||
      466 TGTTCCTCTCTTTTGGGTGCTTATGAGATGGGTTTCTGAGGCCATGAGCCACACA 525
      |||
QY      161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnIleTyrMetCysAsp 180
      |||
      526 GGAAGATTAATGAACCTGACCTTCTGTGCTGACAACTTGTCAATCATTTCAATGTGTGAC 585
      |||
QY      181 IleLeuProLeuGlnLeuLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
      |||
      586 ATCTCTCTCTCTCTGAGCTCTCTGACAGCTTTCATCATGATGAGCTGGTCTTTT 645
      |||
QY      201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
      |||
      646 ATTTGTGTGGCTGTGACGTGAGATGCCATTTGTCACTGCTTTATTTCTTATGCCCC 705
      |||
QY      221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
      |||
      706 ATCTCTCCAGCATTTTACACACAGTTCTTACAGAGCGAGGTCCAAAGCTTTAGACT 765
      |||
QY      241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
      |||
      766 TGCAGTCCCAATTAATGATTTCTCTTTCTTTGTTCTGTGTTCTTATCATGTATCTC 825
      |||
QY      261 LysProProSerIleLeuProLeuaspGlnGlyValValSerSerLeuPheThrThr 280
      |||
      826 AAACCCCTTTCATCTGCCCCCGAGCAAGGAAAGTCTCCCTGTTCTTATACATA 885
      |||
QY      281 ValValProMetPheAsnProLeuIleTySerIleuArgAsnLysAspValLysLeuAla 300
      |||
      886 ATAGTCCCCGTGTAAACCATTAATCTATAGCTTGAGGACAGAGATGTCAAGTTGCC 945
      |||
QY      301 LeuLysArgThrPheSerArgIleSerPheSer 311
      |||
      946 CTGAGGAGAACTTTGGCGAGAAAATCTTTCT 978
      |||
RESULT 15
AX646099      AX646099      1333 bp      DNA      linear      PAT 04-MAR-2003
LOCUS      Sequence 291 from Patent EP1270724.
ACCESSION      AX646099
VERSION      AX646099.1
KEYWORDS      GI:28798478
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE      1.
AUTHORS      Suwa, M., Asai, K., Akiyama, Y. and Aizuratsani, H.
      Guanosine triphosphate-binding protein coupled receptors
      Patent: EP 1270724-A 291 02-JAN-2003;
      National Institute of Advanced Industrial Science and Technology
      (NIA) / Center for Advanced Science and Technology Incubation, Ltd.
      (CPI)
FEATURES
Source      1..1333
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"
      201..1133
      /note="unassigned protein product"
      /codon_start=1
      /protein_id="CAD69184.1"
      /db_xref="GI:28798479"
      /db_xref="EMBL:CAD69184"
      /translation="MAKNSVTEFLIEGLTTPGGRIRIPFLIFLGYTVTVNGNL
      ITLIVNSHLPMTYFLLPMLSLIDFSTTIRPKMLSVSKNLSIFGCTQLTF
      PCFVAVSESLTAMAVADRYVALCNPLLYVTMSCOVCLLLLGAGMGAGMAATG
      SINLTFCAADNLVNHFWCDLIPLEBSNCSYNNELVLFVAVDVGMPLIVTFISYA

```



## ORIGIN

LIISLHNSSTBGRSKAFSTCSHIIIVSLFPGSGAFMYLKLPLSLPLEGRKVSLSL  
YTIIVPLNPLIVSLRNDVAVLRRLTRKIFS"

## Alignment Scores:

Pred. No.:	3,1e-88	Length:	1333
Score:	1273.50	Matches:	250
Percent Similarity:	87.46%	Conservative:	22
Best Local Similarity:	80.39%	Mismatches:	38
Query Match:	80.75%	Indels:	1
DB:	6	Gaps:	1

US-10-023-597-24 (1-311) x AX646099 (1-1333)

```

QY      1  MetaIaIaGIuAsnSerSerSerValThrGluPheIleLeuAlaGIleuAlaHisGln 20
DB      201  ATGGAGGCCAAAAC---TCTTGTGTGACAGAGTTTATCTCGAAGGCTTAAACCCACAG 257
QY      21  ProGIleuGlnValProValPhePheLeuPheLeuGIleuGIleuValAlaValAla 40
DB      258  CCGGACCTGGGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 317
QY      41  GIleuLeuGIleuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
DB      318  GGGAACTCGGCTTGATACCTGATGGCTGAAGTCTCAGCTGCACTCCATCCATGTAC 377
QY      61  PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProIys 80
DB      378  TTCTTCTTTTAACTCTCTTAAATTTCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 437
QY      81  MetLeuMetSerPheValSerArgIleAsnIleIleSerPheThrGIleuGIleuSerGln 100
DB      438  ATGCTGATGATTTTGTCTCAAGGAAGACATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 497
QY      101  PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
DB      498  CTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 557
QY      121  AsparGIleuValGIleuCysAsnProLeuLeuThrIleThrMetSerProGlnVal 140
DB      558  GACCGCTGCGGCTGATGTACCATGTTGTACAGATGACATGCTCTGCGCAGTG 617
QY      141  CysLeuLeuLeuLeuLeuGIleuValIleGIleuValIleGIleuValIleValAlaHisThr 160
DB      618  TGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 677
QY      161  GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisIleIleMetCysAsp 180
DB      678  GGAAACATATGATCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 737
QY      181  ILeuProLeuLeuGIleuSerCysAsnGIleuSerThrIleAsnValIleValIlePhe 200
DB      738  ATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 797
QY      201  ILeuValIleThrValGIleuValIleGIleuValIleValAlaValPheIleSerThrGIleu 220
DB      798  ATTGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 857
QY      221  ILeuSerSerIleLeuArgValSerSerAlaGIleuGIleuArgSerIleValAspPheSer 240
DB      858  ATCTCTCTCAAGCATTTCTACACAGATTTCTACAGAGCAGCTCCAAAGCTTTAGTACT 917
QY      241  CysSerSerThrIleIleIleValSerIleuPhePheGIleuGIleuValAlaPheThrThrLeu 260
DB      918  TGCAGTTCACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 977
QY      261  LysProProSerIleLeuProLeuAspGlnIleValSerSerLeuPheThrThrThr 280
DB      978  AAACCCCTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1037
QY      281  ValValIleProMetPheAsnProLeuIleThrSerLeuArgAsnIleValIleValIleVal 300
DB      1038  ATAGTCCCGGTGTTAAACCATTAATCTATGCTTGAAGAACAGATGTCAAAGTTGCC 1097

```

QY 301 LeuIleArgThrPheSerArgIleSerPheSer 311  
 DB 1098 CTGAGAGAACTTTGGCGGAAAAATCTTTCT 1130

Search completed: September 30, 2004, 11:26:33  
 Job time : 4638 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 08:14:39 ; Search time 447 Seconds

(without alignments)  
2955.681 Million cell updates/sec

Title: US-10-023-597-24

Perfect score: 1577

Sequence: 1 MAENSSSVTEFIAGLIHQ.....LRNDVKALKRTFSRISFS 311

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+pn.model -DEV=xlh  
-Q=cn2.1/USPTO\_SPOOL/US10023597/runat\_29092004\_163409\_14869/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_29Jan04 -QFMT=fastbp -SUFFIX=ring -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10023597\_@CGN\_1\_1\_470\_@runat\_29092004\_163409\_14869 -NCPU=6 -ICPU=3  
-NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_29Jan04:\*  
1: geneeqn19806:\*  
2: geneeqn19906:\*  
3: geneeqn20006:\*  
4: geneeqn2001as:\*  
5: geneeqn2001bs:\*  
6: geneeqn2002s:\*  
7: geneeqn2003as:\*  
8: geneeqn2003bs:\*  
9: geneeqn2003cs:\*  
10: geneeqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1577	100.0	963	6	AAf88413 Human GPC
2	1568	99.4	994	4	AAh31675 Human Olf
3	1515	96.1	960	9	ADD12748 CDNA enco
4	1333.5	84.6	935	4	AAh31668 Human Olf
5	1273.5	80.8	930	4	AAh32421 Human Olf
6	1273.5	80.8	933	5	AAh315910 DNA enco
7	1273.5	80.8	933	5	AAh42233 Human CDN
8	1273.5	80.8	933	6	ABz43067 Human GPC

9	1273.5	80.8	933	6	ABK68455	ABK68455 Human DNA
10	1273.5	80.8	933	6	ABK37521	ABK37521 DNA enco
11	1273.5	80.8	991	6	ABSS58782	ABSS58782 Human G-P
12	1273.5	80.8	1333	9	ADCS5838	ADCS5838 Human GPC
13	1268	80.4	976	6	ABK97207	ABK97207 Human G-P
14	1268	80.4	976	9	AADE0328	AADE0328 Human G-P
15	1266.5	80.3	953	6	ABT05655	ABT05655 GPCR 4 pr
16	1221	77.4	931	9	ADCS6616	ADCS6616 Human GPC
17	1220	77.4	1655	6	AAH24449	AAH24449 Human G-P
18	1180	74.8	933	4	AAH32246	AAH32246 Human Olf
19	1180	74.8	936	5	AAH15911	AAH15911 DNA enco
20	1180	74.8	936	6	ABK40195	ABK40195 Human G-P
21	1180	74.8	942	6	ABZ43068	ABZ43068 Human GPC
22	1180	74.8	951	6	ABK68456	ABK68456 Human DNA
23	1180	74.8	989	6	ABSS5834	ABSS5834 Human G-P
24	1180	74.8	1336	9	ADCS6200	ADCS6200 Human GPC
25	1073	68.0	942	7	ADCS7877	ADCS7877 Human G-P
26	1063	67.4	952	6	ADD12784	ADD12784 CDNA enco
27	1058	67.1	952	6	AAf88431	AAf88431 Human GPC
28	1048	66.5	961	6	ABQ8087	ABQ8087 Human GPC
29	1035	65.6	962	9	ADC79403	ADC79403 Human G-P
30	1024.5	65.0	934	6	AAf88464	AAf88464 Human GPC
31	1023.5	64.9	945	7	ABZ77909	ABZ77909 Human G-P
32	1023	64.9	981	6	AAf88442	AAf88442 Human GPC
33	1018.5	64.6	934	9	ADD12850	ADD12850 CDNA enco
34	1008.5	64.0	984	4	AAH32275	AAH32275 Human Olf
35	1007	63.9	981	9	ADD12806	ADD12806 CDNA enco
36	1006	63.8	950	6	AAf88445	AAf88445 Human GPC
37	1006	63.8	953	6	AAf88433	AAf88433 Human GPC
38	1001.5	63.5	937	6	ABSS58757	ABSS58757 Human G-P
39	1001.5	63.5	937	6	ABSS58759	ABSS58759 Human G-P
40	1001.5	63.5	937	6	ABSS58758	ABSS58758 Human G-P
41	1000	63.4	950	9	ADD12812	ADD12812 CDNA enco
42	999.5	63.4	942	5	AAH42218	AAH42218 Human CDN
43	999.5	63.4	942	6	ABK37508	ABK37508 DNA enco
44	999.5	63.4	1069	3	AAH38958	AAH38958 Human G-P
45	999.5	63.4	1342	9	ADCS5834	ADCS5834 Human GPC

#### ALIGNMENTS

RESULT 1  
AAf88413  
ID AAf88413 standard; CDNA; 963 BP.

AC AAf88413;

DT 12-NOV-2002 (first entry)

DE Human GPCR CDNA SEQ ID 23.

XX Human: anti-HIV; nootropic; antispasmodic; antiarteriosclerotic; GPCR;  
XX immunosuppressive; immunomodulator; cytoskeletal; antiinflammatory; AIDS;  
XX antidiabetic; neuroprotective; anorectic; haemostatic; antibacterial;  
XX fungicide; protozoal; virucide; human G-protein coupled receptor;  
XX gene therapy; vaccine; cardiomyopathy; atherosclerosis; diabetes;  
XX cell signal processing; cancer; obesity; neurodegenerative disorder;  
XX cachexia; anorexia; Alzheimer's disease; Parkinson's disease;  
XX immune disorder; graft versus host disease; bronchial asthma;  
XX Crohn's disease; multiple sclerosis; haemophilia; infectious disease;  
XX idiopathic thrombocytopenic purpura; gene; ss.  
XX Homo sapiens.  
XX OS  
XX WC020250275-A2.  
XX  
XX 27-JUN-2002.  
XX  
XX 18-DEC-2001; 2001MO-US048958.  
XX PF  
XX 18-DEC-2000; 2000US-0256635P.  
XX PR 21-DEC-2000; 2000US-0257876P.  
XX 04-JAN-2001; 2001US-0259743P.

PR 10-JAN-2001; 2001US-0260718P.  
 PR 12-JAN-2001; 2001US-0261498P.  
 PR 24-JAN-2001; 2001US-0263689P.  
 PR 08-FEB-2001; 2001US-0267464P.  
 PR 22-FEB-2001; 2001US-0271021P.  
 PR 14-MAR-2001; 2001US-0275946P.  
 PR 23-MAR-2001; 2001US-0278150P.  
 PR 18-APR-2001; 2001US-0284591P.  
 PR 23-APR-2001; 2001US-0285718P.  
 PR 19-JUN-2001; 2001US-0299327P.  
 PR 16-AUG-2001; 2001US-0312902P.  
 PA (CURA-) CURAGEN CORP.  
 XX  
 PI Padigaru M, Kekuda R, Li L, Ballinger RA, Casman SJ, Spytek KA,  
 PI Baumgartner JC, Burgess CE;  
 XX WPI; 2002-657419/70.  
 DR P-PSDB; AAB71174.  
 XX  
 PT New G-protein coupled receptor polypeptides, useful in gene therapy,  
 PT particularly for treating or preventing cardiomyopathy, atherosclerosis,  
 PT diabetes, Crohn's disease, hemophilia or cancer in humans.  
 XX  
 PS Claim 3; Page 90; 155pp; English.  
 XX  
 CC This invention describes novel human G-protein coupled receptor (GPCRX)  
 CC polypeptides which have anti-HIV, neurotropic, antidiabetic, protozoal,  
 CC antiarteriosclerotic, immunosuppressive, immunomodulator, cytostatic,  
 CC antiinflammatory, antidiabetic, neuroprotective, anorectic, haemostatic,  
 CC antitumor, fungicide and virucide activity. The products of the  
 CC invention can be used in gene therapy or for vaccines. The GPCRX  
 CC polypeptide, GPCRX nucleic acid and antibody are useful for treating,  
 CC preventing or alleviating a GPCRX-associated disorder or a pathological  
 CC state in a subject e.g. cardiomyopathy, atherosclerosis, diabetes, or a  
 CC disorder related to cell signal processing and metabolic pathway  
 CC modulation. The GPCRX polypeptide and nucleic acid are also useful for  
 CC diagnosing the presence of or predisposition to a disease associated with  
 CC altered levels of GPCRX, particularly cancer. These polypeptides, nucleic  
 CC acids and antibodies are also useful for treating or preventing obesity,  
 CC neurodegenerative disorders, AIDS, cancer-associated cachexia, anorexia,  
 CC Alzheimer's disease, Parkinson's disease, immune disorders, graft versus  
 CC host disease, bronchial asthma, Crohn's disease, multiple sclerosis,  
 CC haemophilia, idiopathic thrombocytopenic purpura or infectious disease.  
 CC They can also be used to screen for potential agonist and antagonist  
 CC compounds. The polypeptides are also useful as immunogens to produce  
 CC antibodies or as vaccines. Anti-GPCRX antibodies can be used  
 CC diagnostically to monitor protein levels in tissue as part of a clinical  
 CC testing procedure such as in determining the efficacy of a given  
 CC treatment regimen. The host cells are useful in producing non-human  
 CC transgenic animals which are useful for studying the function and/or  
 CC activity of GPCRX protein and for identifying and/or evaluating  
 CC modulators of GPCRX protein activity. AAF89402-AAF89465 encode the human  
 CC GPCRX proteins represented in AAB71163-AAB71226  
 XX  
 SQ Sequence 963 BP; 193 A; 247 C; 204 G; 319 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 4,12e-137 Length: 963  
 Score: 1577.00 Matches: 311  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Gaps: 0  
 DB: 6

US-10-023-597-24 (1-311) x AAF89413 (1-963)

QY 1 MetAlaIaGluAsnSerSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20  
 DB 2 ATGGCTGCGAAGACTCTCTCTCGTACAGAGTTTATCTCGAGGCTTAATCCACAG 61  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

DB 62 CCGGACATCCAGGATCCCGCTCTTCTCTCTGTTTCTAGTTCCTACGCGGTACGGTGGATG 121  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 DB 122 GGAACCTGGGCTTGATATCTGATAGGCTCAACTCTCCGCTGACATATCCCATGTAC 181  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProlys 80  
 DB 182 TTTTTCCTCCCTCAACTGTCCTGTATATTTAGTTTCTCTACGACCATATATCCCAA 241  
 QY 81 MetLeuMetSerPheValSerArgIleAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 242 ATGCTGATGAGTTTGTCTCAAGAGAAACATTATTTCTTCCACAGGGGTATAGTGTAG 301  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120  
 DB 302 TTCTTCTTCTTCTGTTTCTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 361  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140  
 DB 362 GACCGCTACGGGATGTCATGTAACCCACTGTTGTACAGATCACCATGTCCTCCAGGTG 421  
 QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 DB 422 TGTTCCTCTCTTCTTCTGCTGCTACGCGATGGGGGTTTGGGGCTGTGCTCATACA 481  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 DB 482 GGAATATATATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 541  
 QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 DB 542 ATCTTCTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 601  
 QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 DB 602 ATGTTGTACCGTGGATTTGGGTGCCATGTTGCCGTTTATCTTCTTATGCTTTT 661  
 QY 221 IleLeuSerSerIleLeuArgValSerSerArgIleGlyIleGlySerTyrAlaPheSerSer 240  
 DB 662 ATCTTCTCAGATTTCTCGGGTATGTTCTGTCAGGGCAGAGCTTAAAGCTTCAGTAGC 721  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260  
 DB 722 TGACACTCTCTCATATATGACAGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 781  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThr 280  
 DB 782 AAACCCCTTCCATTTTACCCCTGGACCAAGGGAAGGTCTCTCTCTTCTTATACCACT 841  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300  
 DB 842 GTGGTGCCCATGTTTAAACCATTAATCAACAGCTGAGAGATAAGATGTCMAAATTGCC 901  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 DB 902 CTGAAGAGAACCTTTTCCAGATTAAGCTTTTCT 934

## RESULT 2

AAH31675 standard; DNA; 994 BP.

AAH31675;

30-JUL-2001 (first entry)

Human olfactory receptor polynucleotide, SEQ ID NO: 248.

Human; olfactory receptor; OR; primary scent determination;

secondary scent determination; polypeptide library; odour receptor;

scent profile; scent fingerprint; scent representation; ds.

Homo sapiens.

XX WO200127158-A2.  
 PN 19-APR-2001.  
 XX 06-OCT-2000; 2000WO-US027582.  
 XX 08-OCT-1999; 99US-0158615P.  
 PR 24-FEB-2000; 2000US-0184809P.  
 XX (DIGI-) DIGISCENTS.  
 PA (YEDA ) YEDA RES & DEV CO LTD.  
 PI Bellensen J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
 XX WPI; 2001-290713/30.  
 DR New polynucleotides which encode polypeptides involved in olfactory  
 PT sensation for identifying olfactory agonists and antagonists.  
 XX Claim 8; Page 276; 1857pp; English.  
 XX The present sequence is one of a number of isolated polynucleotides which  
 CC encode polypeptides involved in olfactory sensation. The polynucleotides  
 CC can be used in screening for olfactory agonists and antagonists. The  
 CC methods allow for the determination of primary scents and the  
 CC identification of the odor receptors used to detect these primary  
 CC scents. The methods also enable determination of secondary scents and the  
 CC identification of combinations of odor receptors that are involved in  
 CC detecting such secondary scents. This enables the construction of a scent  
 CC representation (also called a scent fingerprint or scent profile), which  
 CC may be used to re-create and edit scents. Libraries of olfactory  
 CC receptors are useful for determining the interaction pattern of a  
 CC composition with the receptors, and can be used for determining  
 CC differences in the olfactory faculties of different individuals  
 CC  
 XX Sequence 994 BP; 201 A; 250 C; 207 G; 336 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2, 93e-136 Length: 994  
 Score: 1568.00 Matches: 310  
 Percent Similarity: 99.68% Conservative: 0  
 Best Local Similarity: 99.68% Mismatches: 1  
 Query Match: 99.43% Indels: 0  
 DB: Gaps: 0  
 US-10-023-597-24 (1-311) x AAH31675 (1-994)  
 QY 1 MetAlAlAGluAnsSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20  
 DB 1 ATGGCTGCCGAGAACCTCTCTCCGACAGAGTTATCTCGCAGGCTTAATCCACACG 60  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40  
 DB 61 CCGGAGCTCCAGGATCCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAnsSerArgLeuHisIleProValTyr 60  
 DB 121 GGGACCTGGGCTTGATTAATCTGATAGGCTCACTCTCGCTGCATATCCCAATGTAC 180  
 QY 61 PhePheProPheLeuLeuSerLeuValAspPheSerPheSerThrThrIleLeuProLys 80  
 DB 181 TTTTTCCTTCACTGCTGCTCTGATAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240  
 QY 81 MetLeuSerSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 241 ATGCTGATGAGTTTGTCTCAAGAGAACATATTCTCTCTCAAGGATGATGAGTACG 300  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120  
 DB 301 TTCT 360  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

DB 361 GACCGCTACCTGGGACATCTGTAAACCACTGTTTATACCATCACTGCTCCCGAGT 420  
 QY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 DB 421 TGTGTCTCTTCTTACCTGGGTGTCTACGGGATGGGGATTTTGGGGCTGTGGCTCATACA 480  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 DB 481 GGAATATATAGTGTCTTACCTCTTGTGACAGACACCTGTCAATCATCATATGTGTAC 540  
 QY 181 IleuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 DB 541 ATCTTCCCTTCTTGTAGCTCTCTGTAAGGCTCTTACATTAATATGCTGTGATCTTT 600  
 QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 DB 601 ATTGTTGTACCGTTGGCATTTGGGTGTCCTCATTTGTGCGTTTATCTCTTAAGTTT 660  
 QY 221 IleuSerSerIleLeuArgValSerSerAlaGlyGlyArgSerIleAlaPheSerSer 240  
 DB 661 ATTCTTCCAGCATCTCTCCGCTTACTGCTGAGGGCAGGCTTAAGCTTCAGTAGC 720  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260  
 DB 721 TGCACTCTCAATTAATGGAGTTTCTTTTCTTTGGGTCAAGAGCTTTTACGTACTTC 780  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280  
 DB 781 AAACCCCTTCATTTACCTCGACCAAGGAGAAATGCTCTCTCTCTATACCACT 840  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnValLysValLysLeuAla 300  
 DB 841 GTGGTCCCATGTTTAAACCATTAATCTACAGCTGAGATTAAGATGCAAACTTGCC 900  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 DB 901 CTGAAGAGAACCTTTTCCAGATTAAGCTTTTCT 933  
 RESULT 3  
 ID ADD12748 standard; cDNA; 960 BP.  
 AC ADD12748;  
 XX 01-JAN-2004 (first entry)  
 DE cDNA encoding novel human olfactory receptor (OR) seq id 12.  
 XX cardiant; antiarteriosclerotic; antidiabetic;  
 XX G-Protein Coupled Receptor modulator; gene therapy; olfactory receptor;  
 XX G-Protein Coupled Receptor X-associated disorder;  
 XX GPCR associated disorder; cardiomyopathy; atherosclerosis;  
 XX cell signal processing; metabolic pathway modulation; cancer; diabetes;  
 XX human; G-protein coupled receptor; GPCR; gene; ss.  
 OS Homo sapiens.  
 PN US2003109692-A1.  
 PD 12-JUN-2003.  
 XX 18-DEC-2001; 2001US-00023597.  
 XX 18-DEC-2001; 2001US-0256635P.  
 XX 18-DEC-2000; 2000US-0257876P.  
 XX 21-DEC-2000; 2000US-0257876P.  
 XX 04-JAN-2001; 2001US-0259743P.  
 XX 10-JAN-2001; 2001US-0260718P.  
 XX 12-JAN-2001; 2001US-0261498P.  
 XX 24-JAN-2001; 2001US-0261698P.  
 XX 08-FEB-2001; 2001US-0267464P.  
 XX 22-FEB-2001; 2001US-0271021P.  
 XX 14-MAR-2001; 2001US-0275946P.



PF 06-OCT-2000; 2000WO-US027582.  
 XX 08-OCT-1999; 99US-0158615P.  
 PR 24-FEB-2000; 2000US-0184809P.  
 XX (DIGI-) DIGISCENTS.  
 PA (YEDA) YEDA RES & DEV CO LTD.  
 PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
 DR WPI, 2001-290713/30.  
 XX  
 PT New polynucleotides which encode polypeptides involved in olfactory  
 PT sensation for identifying olfactory agonists and antagonists.  
 XX  
 PS Claim 8; Page 273; 1857pp; English.  
 XX  
 CC The present sequence is one of a number of isolated polynucleotides which  
 CC encode polypeptides involved in olfactory sensation. The polynucleotides  
 CC can be used in screening for olfactory agonists and antagonists. The  
 CC methods allow for the determination of primary scents and the  
 CC identification of the odour receptors used to detect these primary  
 CC scents. The methods also enable determination of secondary scents and the  
 CC identification of combinations of odour receptors that are involved in  
 CC detecting such secondary scents. This enables the construction of a scent  
 CC representation (also called a scent fingerprint or scent profile), which  
 CC may be used to re-create and edit scents. Libraries of olfactory  
 CC receptors are useful for determining the interaction pattern of a  
 CC composition with the receptors, and can be used for determining  
 CC differences in the olfactory faculties of different individuals

XX SQ Sequence 935 BP, 189 A, 238 C, 190 G, 318 T, 0 U, 0 Other;

#### Alignment Scores:

Pred. No.:	1,63e-114	Length:	935
Score:	1333.50	Matches:	257
Percent Similarity:	92.18%	Conservative:	26
Best Local Similarity:	83.71%	Mismatches:	23
Query Match:	84.56%	Indels:	2
DB:	4	Gaps:	1

US-10-023-597-24 (1-311) x AAH31668 (1-935)

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20  
 DB 1 ATGCTGCGCAAGAACTCTCTCCGTGACAGCGTTTATCTTCAGCTTAACCGACCG 60  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrVal 40  
 DB 61 CCGGAGACTCCAGATCCCGCTTCTCTCTGTTTCTTAAGTTTCAACGGGTCACGGTGGTG 120  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIlePheMetTyr 60  
 DB 121 GGGAACTCGGCTTGATATCTGATAGGGCTCAACTCGCTGCAATATCCCATTTAC 180  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80  
 DB 181 TTTTCCCTTCAACTGCTCTCATAGATTGATTATTCACACTACCCCTCGCCCTTAA 240  
 QY 81 MetLeuMetSerPheValSerValAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 241 ATGCTGATGACCTTGTCTCA--GAGAACATCAATTCCTATGACGAGGTATGACTCAG 297  
 QY 101 PhePhePhePheCysPhePheValPheSerGlnSerPheIleLeuSerAlaMetValGln 120  
 DB 298 CTTTCTTCTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 357  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIlePheMetSerProGlnVal 140  
 DB 358 GACCGCTTCGTGGGATCTGTACCCCACTGTGTACCGGTCACCACTGTCTCCCGAGATG 417  
 QY 141 CysIleuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160

DB 418 TGTTCCTCTTACTGCGGTGTCATGAGGAGAT-TTGGGGCTGTGCTCATATG 476  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnIleTyrMetCysAsp 180  
 DB 477 GGAACATATATGTTATGCTCTTTTGTGAGACACACTTGTCAATCATATATGTTGAC 536  
 QY 181 IleuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 DB 537 ATTCCTCTCTCTGAGCTCTCTGCAACAGCTCTTCAATATATTTGTGTGGTGT 596  
 QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 DB 597 ATTATGTGACCGTTGGCATTTGGGGTCCGATTTGTACACATTTTCTCTTATGTTT 656  
 QY 221 IleuSerSerIleLeuArgValSerSerAlaGlnGlyArgSerTyrAlaPheSerSer 240  
 DB 657 ATTCCTTCCAGCATTCCTCAACATTAGTCCACAGAGGAGGAGCTTAAACCTTCACTAC 716  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260  
 DB 717 TGCAGTCCCAATATGTTATGCTTCTTTTGGGTCAAGTCTTTCATATGTTACCTC 776  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThr 280  
 DB 777 AAACCACTCTTATTTCTTACCCCTGACCAAGGAGAAAGTCTTCTTCTTGTACTGCT 836  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnIleValIleValAla 300  
 DB 837 GTGGTCCCATGTTTATCCATTTATCTTACCTGACGCTGAGAAATTAATGTTCAAGTTGC 896  
 QY 301 LeuLysArgThrPheSerArg 307  
 DB 897 CTGAGAGAACTTTTTCAGA 917

#### RESULT 5

ID AAH32421 standard; DNA, 930 BP.

XX AAH32421;

DT 30-JUN-2001 (first entry)

DE Human olfactory receptor polynucleotide, SEQ ID NO: 994.

KX Human; olfactory receptor; OR; primary scent determination;

KW secondary scent determination; polypeptide library; odour receptor;

OS scent profile; scent fingerprint; scent representation; de.

XX Homo sapiens.

XX W0200127158-A2.

PD 19-APR-2001.

PF 06-OCT-2000; 2000WO-US027582.

XX 08-OCT-1999; 99US-0158615P.

PR 24-FEB-2000; 2000US-0184809P.

XX (DIGI-) DIGISCENTS.

PA (YEDA) YEDA RES & DEV CO LTD.

PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;

DR WPI, 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory

PT sensation for identifying olfactory agonists and antagonists.

XX Claim 8; Page 577; 1857pp; English.

CC The present sequence is one of a number of isolated polynucleotides which

CC encode polypeptides involved in olfactory sensation. The polynucleotides

CC can be used in screening for olfactory agonists and antagonists. The  
 CC methods allow for the determination of primary scents and the  
 CC identification of the odor receptors used to detect these primary  
 CC scents. The methods also enable determination of secondary scents and the  
 CC identification of combinations of odor receptors that are involved in  
 CC detecting such secondary scents. This enables the construction of a scent  
 CC representation (also called a scent fingerprint or scent profile), which  
 CC may be used to re-create and edit scents. Libraries of olfactory  
 CC receptors are useful for determining the interaction pattern of a  
 CC composition with the receptors, and can be used for determining  
 CC differences in the olfactory faculties of different individuals  
 XX  
 SO Sequence 930 BP; 190 A; 251 C; 188 G; 301 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 6,07e-109 Length: 930  
 Score: 1273.50 Matches: 250  
 Percent Similarity: 87.46% Conservative: 22  
 Best Local Similarity: 80.39% Mismatches: 38  
 Query Match: 80.75% Indels: 1  
 Gaps: 1

US-10-023-597-24 (1-311) x AAH32421 (1-930)

QY 1 MetAlaAlaGluAnSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20  
 Db 1 ATGGACGCCAAAAAC---TCTTCTGACAGATTATCTCGAAGGCTTAACCCACAG 57  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheThrAlaValThrValAl 40  
 Db 58 CCGGACATGCGACCCCT 117  
 QY 41 GlyAsnLeuGlyLeuIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 Db 118 GGGACCTGGCTTGAATTAACCTGATGGCTGAACCTCACCTGACACCTCCATGTAC 177  
 QY 61 PhePheProPheLeuSerLeuValAspPheSerPheSerThrIleLeuProlys 80  
 Db 178 TTTCTCTTTTAACTCTCTTTAATGATTTCTGTTTCCACATCACTATCTCCGMAA 237  
 QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100  
 Db 238 ATGCTGATGAGTTTGTCTCAAGAGAACATCATTTCTTCCAGGGGTATGACTCAG 297  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120  
 Db 298 CT 357  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140  
 Db 358 GACCGCTACGTCGACATCTGTAACTGTTGTAACAGTCATGTCATGTCGACAGG 417  
 QY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 Db 418 TGTTCCTCTCTTGTGTCCTTGTGTCCTTGTGTCCTTGTGTCCTTGTGTCCTTGTG 477  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 Db 478 GGAAGCATATGAACCTGACCTTCTGTCGACCACTGTCATCATTCATTCATGTCAC 537  
 QY 181 IleLeuProLeuLeuLeuLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 Db 538 ATCTTCCTCTCTTGTAGCTCTCTGACACCTCTTACATGAAGACCTGATGCTTT 597  
 QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 Db 598 ATTGTGTGCTGCTTGAATGGAATGCCATTCATGTCATGCTTATTTCTTATGCCCTC 657  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGlyGlyArgSerTyrAlaPheSerSer 240  
 Db 658 ATCTCTCCAGCATTTACACAAAGTTTACAGGAAGGCTCCAAAGCCCTTAGACT 717  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260

Db 718 TCCAGTCCCAATATGATGTTCTCTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 777  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThr 280  
 Db 778 AAACCCCTTTCATCTCTGCTCCCTGACGAGAAAGTCTCTCTCTCTTATACATA 837  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValValLeuAla 300  
 Db 838 AATGTCCTCCGTGTAAACCCATTTATCTATGTTGAGCAACAAAGATGTCAAAGTTGCC 897  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 Db 898 CTGAGGAGAACTTTGGCGAAGAAATCTTTCT 930

RESULT 6  
 AAS15910  
 ID AAS15910 standard; cDNA, 933 BP.

XX AAS15910;  
 AC

DT 25-JUN-2002 (first entry)

DE DNA encoding G-protein coupled receptor (GCREC) #14.

XX G-protein coupled receptor; GCREC; vaccine; gene therapy;  
 KM cell proliferation disorder; cancer; arteriosclerosis;  
 KM neurological disorder; epilepsy; stroke; cardiovascular disorder;  
 KM hypertension; ischaemic heart disease; gastrointestinal disorder;  
 KM anorexia; peptic ulcer; autoimmune disorder; inflammatory disorder;  
 KM diabetes; osteoporosis; psoriasis; metabolic disorder; obesity;  
 KM schizophrenic disorder; neuroskeletal disorder; ss.

XX Homo sapiens.

OS Key Location/Qualifiers

FT CDS 1..933

FT /tag= a

FT /product= "GCREC 14"

FT /note= "G-protein coupled receptor 14"

PN MO200166742-A2.

XX 13-SEP-2001.

PD 01-MAR-2001; 2001MO-US006814.

XX 03-MAR-2000; 2000US-0186854P.

PR 10-MAR-2000; 2000US-0186834P.

PR 17-MAR-2000; 2000US-0190453P.

PR 20-MAR-2000; 2000US-0190730P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lal P, Tang YT, Patterson C, Yao MG, Shih LL, Tribouley CM;  
 PI Lu DM, Yue H, Khan FA, Policky JU, Au-Young J, Yang J, Harland L;  
 PI Walsh RT, Lo TP, Borowsky MB;

XX WPI; 2001-656776/75.  
 DR P-PSDB; AAI0313.

XX Novel G-protein coupled receptor polypeptides, for treating and  
 PT preventing autoimmune/inflammatory disorders, neurological disorders,  
 PT cell proliferative disorders, cardiovascular disorders and viral  
 PT infections.

XX Claim 5; Page 138; 141pp; English.

CC The invention describes a novel isolated polypeptide, selected from a  
 CC group of 21-G-protein coupled receptor polypeptides (GCREC) and useful in  
 CC vaccines and gene therapy. The polypeptide (I) is useful for screening  
 CC for agonist or antagonist of (I), compounds specifically binding to (I),  
 CC or compounds that modulate the activity of (I). The polynucleotide



CC encoding (1) is useful for screening a compound for effectiveness in  
 CC altering expression of a target polynucleotide comprising (ii), by  
 CC exposing a sample comprising the target polynucleotide to a compound,  
 CC detecting altered expression of the target polynucleotide, and comparing  
 CC the expression of the target polynucleotide in the presence of varying  
 CC amounts of compound and in the absence of the compound. (i) and (ii) are  
 CC useful for diagnosis, treatment and prevention of cell proliferative  
 CC disorders (e.g. cancers, aretioclerosis, atherosclerosis), neurological  
 CC disorders (e.g. epilepsy, stroke, schizophrenic disorders and  
 CC neuroskeletal disorders), cardiovascular disorders (e.g. hypertension,  
 CC ischaemic heart disease), gastrointestinal disorders (e.g. anorexia,  
 CC peptic ulcer), autoimmune/inflammatory disorders (e.g. diabetes mellitus,  
 CC osteoporosis, psoriasis), and metabolic disorders such as obesity.  
 CC Furthermore, the polynucleotide is useful, as primers for detecting  
 CC single nucleotide polymorphisms, as elements in microarray, to monitor or  
 CC measure protein-protein interactions, drug-target interactions, and gene  
 CC expression profiles; to generate a transcript image of a tissue or cell  
 CC type, and to generate hybridisation probes useful in mapping the  
 CC naturally occurring genomic sequence. This sequence encodes G-protein  
 CC coupled receptor 14, one of 21 GPCR proteins described in the method of  
 CC the invention

XX  
 XX Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

# Alignment Scores:

Prod. No.:	6,09e-109	Length:	933
Score:	1273.50	Matches:	250
Percent Similarity:	87.46%	Conservative:	22
Best Local Similarity:	80.39%	Mismatches:	38
Query Match:	80.75%	Indels:	1
DB:	5	Gaps:	1

US-10-023-597-24 (1-311) x AAS15910 (1-933)

QY 1 MetAlaAlaGluSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20  
 DB 1 ATGGCACCCAAAAC--TTCTCTGACAGAGTTTCTCCAGAGCTTAACCCACG 57  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValAl 40  
 DB 58 CCGGAGATCGGAGATCCCTCTTCTCTTCTGATTTCTTACACGGTCACCGTGGTG 117  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 DB 118 GGGAACTGGGCTTGATACCCGATGGGCTGATCTCACTGACACTCCCACTGAC 177  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80  
 DB 178 TTCTTCTTTTAACTCTCTTTATATGATTTCTGTTTCCACTACCATCACTCCCAA 237  
 QY 81 MetLeuMetSerPheValSerArgValAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 238 ATGCTGATGATTTGCTCAAGAGAACATCATTTCTTCAAGGCTGATGACTCAG 297  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120  
 DB 298 CTCTTCTTCTTCTCTTCTTCTTCTGCTGAGTCTTCACTGCTGACGAGATGGGAT 357  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140  
 DB 358 GACCGCTACGCGCATCTGTACCACTGTTGTACACAGTCACTCACTGTTCCAGGTG 417  
 QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyAlaIleGlyAlaValAlaHisThr 160  
 DB 418 TGTGTGCTCTTGTGTGGGCTATGAGATGGAGTTGCTGGGCGCATGCGCCACACA 477  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 DB 478 GGAAGCATATGAACTGACTCTGTGCTGACAACTGTGCATTCATTCATGCTGTGAC 537  
 QY 181 IleLeuProLeuLeuGlnLeuSerCysAsnGlySerTyrIleAsnValIleuValIlePhe 200  
 DB 538 ATCTTCTCTCTGAGCTCTCTGCAACAGCTCTTACATGATGAGCTGGTGCTTT 597

QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 DB 598 ATTCGTGGCTGTGTGACGTTGAGATGCCATTTGCTGCTTATTTCTTATGCCCC 657  
 QY 221 IleLeuSerSerIleLeuAlaGValSerSerAlaGlnGlyArgSerTyrAlaPheSerSer 240  
 DB 658 ATCTCTCCAGCATTTTACCAACAGTTTTCACAGAAAGGAGGTCCAAAGCCTTAGTACT 717  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260  
 DB 718 TCGATTTCCCATATATGATTTCTCTTCTTCTGTTCTGAGTTCATGATATCTC 777  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280  
 DB 778 AAACCTTTTCATCTCCCTCTGACCAAGGAAAGTGTCTCCCTGTTCTATACATA 837  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300  
 DB 838 ATAGTCCCGGTTAAACCATTTATCTATGCTTGAGAACAGATGCAAGTCC 897  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 DB 898 CTGAGAGAACTTTGGCAGAAAATCTTTCT 930

RESULT 7  
 AAS42233  
 ID AAS42233 standard; CDNA; 933 BP.  
 AC AAS42233;  
 XX 18-DEC-2001 (first entry)  
 DT  
 DE Human cDNA encoding olfactory receptor AORR26.  
 KW Human; olfactory receptor; G protein-coupled receptor; GPCR; odourant;  
 KW ss; food additive; cosmetic; fragrance; pharmaceutical additive.  
 CS Homo sapiens.  
 OS  
 PN WO200168805-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 13-MAR-2001; 2001WO-US007771.  
 XX  
 PR 13-MAR-2000; 2000US-0188914P.  
 PR 24-MAR-2000; 2000US-0192033P.  
 PR 12-APR-2000; 2000US-0198474P.  
 PR 24-APR-2000; 2000US-0199335P.  
 PR 26-MAY-2000; 2000US-0207702P.  
 PR 23-JUN-2000; 2000US-0213849P.  
 PR 16-AUG-2000; 2000US-0226534P.  
 PR 07-SEP-2000; 2000US-0230732P.  
 PR 07-FEB-2001; 2001US-0266862P.  
 PA (SENO-) SENOMTX INC.  
 XX  
 PI Zozulya S;  
 XX  
 DR WPI; 2001-570867/64.  
 DR P-PSDB; AAT24540.  
 XX  
 PT Nucleic acids encoding human olfactory G protein-coupled receptors,  
 PT useful for screening for compounds involved in olfactory sensation, where  
 PT the compounds can be used in the food, pharmaceutical and cosmetic  
 PT industries to customize odors.  
 XX  
 PS Claim 1; Page 94; 319p; English.  
 XX  
 CC The invention relates to nucleic acids encoding human olfactory  
 CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's  
 CC specifically recognise molecules, odourants, that elicit specific

CC olfactory sensation. The human olfactory receptors and polynucleotides  
 CC encoding them are useful for screening a library of chemical compounds  
 CC for compounds that are involved in olfactory sensation. Modulators of  
 CC their activity are useful for pharmacological and genetic modulation of  
 CC olfactory signaling pathways. Therefore, they can be used in the food,  
 CC pharmaceutical and cosmetic industries to customise odours and  
 CC fragrances. The present sequence encodes a human olfactory receptor of  
 CC the invention

XX  
 SO Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	Length:	933
Score: 1273.50	Matches: 250	
Percent Similarity: 87.46%	Conservative: 22	
Best Local Similarity: 80.39%	Mismatches: 38	
Query Match: 80.75%	Indels: 1	
DB: 5	Gaps: 1	

US-10-023-597-24 (1-311) x AAS42233 (1-933)

```

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleValAlaGlyLeuIleHisGln 20
Db 1 AAGGACGACCAAAAC---TCTTCTGTGACAGATTATCTCGAAGGCTTAACCCACG 57
QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyAlaValThrValAl 40
Db 58 CCGGGGCTGCGGATCCCCCTCTTCTCTGTTTGGGTTTCTACACGGTCACCGTGTG 117
QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
Db 118 GGAAGACCTGGGCTTAACCTGATGGGCTGAACCTCACTCACTCACTCACTCACTGAC 177
QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProTyr 80
Db 178 TCTTCTCTTTTAACCTCTCTTATATAGATTCTGTTCTCTCACTCACTCACTCACTCA 237
QY 81 MetLeuMetSerPheValSerArgLeuAsnIleIleSerPheThrGlyCysMetSerGln 100
Db 238 ATGCTGATGAGTTTGTCTCAAGAGAAACATCTTCTTCAACAGGCTGTATGACTCAG 297
QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
Db 298 CTCTTCTCTTCTGCTCTTCTGCTCTGAGTCTCTTCACTCACTCACTCACTCACTGAT 357
QY 121 AspArgTyrValGlyTleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
Db 358 GACCCCTACGTCGACATCTGTACCACTGTTGTACACAGACACATGTCCTTGCACAGTG 417
QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyValAlaAlaHisThr 160
Db 418 TGTTCCTCTCTTCTTGTGGTCCCTATGGAGGAGGTTTCTGCGGSCCATGGCCACACA 477
QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
Db 478 GGAAGCATATGAACCTGACCTCTGTGCTGACAACTTGTCACTATTCATGTGTGAC 537
QY 181 IleuPheLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIleValIlePhe 200
Db 538 ATCTTCTCTCTGAGCTCTCTCTGCAACGCTTCACTCACTGATGAGCGTGCTTTT 597
QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
Db 598 ATTGGGGGCTGTGAGCTGTGAGTGAATGCCATGTCTACTGTTTATTTTATGCCCTC 657
QY 221 IleuLeuSerIleLeuArgValSerSerAlaGluGlyArgSerIlyAlaPheSerSer 240
Db 658 ATCTCTCCAGCATTTCAACACAGATTCTCAAGAAAGCGGTCACAAAGCTTTAGACT 717
QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
Db 718 TGCAGTCCACATATTTAGTTCTCTTCTTGTGTTCTGTGCTTCACTGATATCTC 777

```

```

QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThr 280
Db 778 AAACCCCTTTCCATCTCCGCCCCGAGCAAGGAGAGTCTCTCCCTGTATACCA 837
QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
Db 838 ATACTCCCTGCTGTAAACCATTAATCTATAGCTTGAAGAACAGAGATGTCAAAGTTC 897
QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311
Db 898 CTGAGAGAACTTTGGGCAAGAAAAATCTTTCT 930

```

# RESULT 8

ABZ43067  
 ID ABZ43067 standard; DNA; 933 BP.  
 XX  
 AC ABZ43067;  
 XX  
 DT 06-MAR-2003 (first entry)  
 XX  
 DE Human GPCR polynucleotide SEQ ID NO 395.

XX  
 KW Human; GPCR; G protein coupled receptor; signal transduction; olfactory;  
 KW drug development; gustatory; taste; fragrance; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200216548-A2.  
 XX

PD 28-FEB-2002.

XX 30-JUL-2001; 2001WO-IB001446.

XX 04-AUG-2000; 2000JP-00237818.

PR 13-FEB-2001; 2001JP-00034434.

XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.

PI Haga T, Takeda S, Mitaku S;  
 XX

DR WPI: 2002-304118/34.

XX P-PSDB; ABP95793.

PT Database global search for G protein-coupled receptors, proteins and  
 PT encoded genes for studying in vivo signal transduction mechanism and  
 PT identifying targets for drug development.  
 XX

XX Claim 9; SEQ ID NO 395; 97pp + Sequence Listing; Japanese.

XX The invention relates to a method for screening G protein-coupled  
 CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-  
 CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane  
 CC domains with 250-1000 amino acid residues to give a gene homologous with  
 CC a known GPCR gene. The receptor proteins and encoded genes are useful for  
 CC studying in vivo signal transduction mechanism and identifying targets  
 CC for drug development e.g. based on olfactory and gustatory receptors in  
 CC form of agonists and antagonists by screening intrinsic and extrinsic  
 CC ligands as bitter taste inhibitors, taste enhancers and fragrance  
 CC improves. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at [ftp.wipo.int/pub/published\\_pcr\\_sequences](http://wipo.int/pub/published_pcr_sequences)  
 XX

SO Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	Length:	933
Score: 1273.50	Matches: 250	
Percent Similarity: 87.46%	Conservative: 22	
Best Local Similarity: 80.39%	Mismatches: 38	
Query Match: 80.75%	Indels: 1	
DB: 6	Gaps: 1	

US-10-023-597-24 (1-311) x ABZ43067 (1-933)

```

QY      1 MetAlaAlaGluAsnSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20
DB      1 ATGGCAGCCAAAAC---TCTTCGTGACAGAGTTATCTCCGAAAGGCTTAACCCACGAG 57
QY      21 ProGlyLeuGluValProValPhePhePheLeuGlyPheValValThrValVal 40
DB      58 CCGGAGCTGGGATCCCTCTTCTTCTGTTCTGTTCTGAGGTTCTTACAGGCTCAGGTTG 117
QY      41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
DB      118 GGAACCTGGCTTGATTAACCTGATGGCTGAAGCTCAGCTCAGACATCCCATGATAC 177
QY      61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleLeuProTyr 80
DB      178 TTCTTCCTTTTAACTCTCTTAAATAGATTCTGTTCTCTCCATCAACATCACTCCCAA 237
QY      81 MetLeuMetSerPheValSerArgIysAsnIleIleSerPheThrGlyCysMetSerGln 100
DB      238 ATGCTGATGAGTTTCTCTCAAGAAAGATCATTTCTTCAAGGGGTATGACTCAG 297
QY      101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
DB      298 CTCCTCTCTCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 357
QY      121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
DB      358 GACCGCTACGCGCCATCTGTAAACCACTGTTGTACACAGTCATGCTTCCAGAGTG 417
QY      141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
DB      418 TGTTCCTCTCTTGTGTTGGGCTGATGGAGTGGGTTGCTGGGAGCCATCCACACA 477
QY      161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
DB      478 GGAAGCATATGAACCTGACCTCTGTGCTGACCACTGCTCAATCATTTCAATGATGAC 537
QY      181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValIleValIlePhe 200
DB      538 ATCTTCCTCTCTCTGAGCTCTCTGCAACAGCTTATATGAAGAGCTGTGCTTT 597
QY      201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
DB      598 ATGTGTGTGCTGTGACCTTGAATGCCATTTGCACGTGTTATTTCTTATGCCCCCTC 657
QY      221 IleLeuSerSerIleLeuArgValSerSerAlaGlnGlyArgSerIysAlaPheSerSer 240
DB      658 ATCTCTTCAGCATTTCTACACACAGCTTCTACAGAAAGAGGTCCAAAGCTTTAGTACT 717
QY      241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
DB      718 TGCAGTTCCACATATTTAGTTTCTCTTTCTTTGTTGTTGTTGCTTTCATGATCTC 777
QY      261 IysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheThrThr 280
DB      778 AAACCCCTTTCATCTGCCCCCTGAGCAAGGAAAGTGTCTCTGTTCTATACCATTA 837
QY      281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnIysAspValIysLeuAla 300
DB      838 ATAGTCCCGGTGTTAAACCATTAATCTAGTTTGAAGAAACAGATGTCAAGTTGCC 897
QY      301 LeuLysArgThrPheSerArgIleSerPheSer 311
DB      898 CTGAGAGAGAACTTTGGCAGAAAATCTTTTCT 930

```

RESULT 9  
ABK68455 standard; DNA; 933 BP.

AC AAK68455;  
AC  
DT 02-JUL-2002 (first entry)  
XX

```

DE      Human DNA for olfactory and pheromone G protein-coupled receptor #55.
XX      Human; db; gene; olfactory and pheromone G protein coupled receptor;
XX      GPCR; trianquillizer; antidepressant; neuroleptic; endocrine; anabolic;
XX      anorectic; taste; fragrance; food additive; cosmetic; cell migration;
XX      sterility; psychotic disorder; neurological disorder; anxiety;
XX      schizophrenia; manic depression; depression; axonal growth;
XX      menstrual cycle; appetite sexual motivation; sexual attraction;
XX      aggression.
XX      Homo sapiens.
XX      OS
XX      WO200224726-A2.
XX      PD
XX      28-MAR-2002.
XX      PF
XX      21-SEP-2001; 2001MO-BE000162.
XX      PR
XX      22-SEP-2000; 2000EP-00870211.
XX      PA
XX      (CHEM-) CHEMCOM SA.
XX      P1
XX      Veithen A;
XX      WPI; 2002-330013/36.
XX      DR
XX      P-PSDB; AAU95568.
XX      PT
XX      Novel pheromone G-protein coupled receptor and receptor-derived agonists,
XX      antagonists or inhibitors useful in food or cosmetic products or in the
XX      treatment or prevention of neurological disorders such as anxiety and
XX      schizophrenia.
XX      PT
XX      Disclosure; Page 178-179; 833p; English.
XX      PS
XX      The invention relates to olfactory and pheromone G-protein coupled
XX      receptor (GPCR) or a protein 95% identical to the GPCR, a specific active
XX      portion and its encoding polynucleotide. Also included are an agonist,
XX      antagonist or inhibitor of the GPCR or the polynucleotide, a vector
XX      comprising the polynucleotide, a cell transformed by the vector, a non-
XX      human mammal comprising a partial or total deletion of the polynucleotide
XX      encoding the receptor and screening (detection and possibly, recovering)
XX      of compounds which are known or not known to be agonist, antagonists or
XX      inhibitors of natural compounds to the GPCR. The receptor-derived
XX      agonists, antagonists, inhibitors or compounds are used as an
XX      improvement, elimination or substitution of an existing taste and/or a
XX      fragrance of (or in) the food and/or cosmetic products. They can also be
XX      used in the preparation of medicament in the treatment and/or prevention
XX      of a mammalian disorder, such as cell migration, sterility, psychotic and
XX      neurological disorders, including anxiety, schizophrenia, manic
XX      depression, depression, for promoting axonal growth, nerve cell
XX      connection and nerve regeneration for modulating male and female
XX      endocrine functions, hormone production and the menstrual cycle, for the
XX      prevention or the treatment by stimulation of several mammalian
XX      behaviour, such as stimulation or suppression of appetite, sexual
XX      motivation, sexual attraction, aggression and for promoting or
XX      suppressing chemical communication between organisms. The present
XX      sequence is a human DNA encoding an olfactory and pheromone GPCR
XX      SQ
XX      Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

```

## Alignment Scores:

Pred. No.:	Length:	933
Score:	1273.50	250
Percent Similarity:	87.46%	22
Best Local Similarity:	80.39%	38
Query Match:	80.75%	1
DB:	6	1

US-10-023-597-24 (1-311) X ABK68455 (1-933)

```

QY      1 MetAlaAlaGluAsnSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20
DB      1 ATGGCAGCCAAAAC---TCTTCGTGACAGAGTTATCTCCGAAAGGCTTAACCCACGAG 57

```

```

QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
DB 58 CCGGACACGCGGATCCCTCTTCTTCTGTTTCTGAGTTCTTCAACCGTCAACCGTGTG 117
QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
DB 118 GGAACCTGGGGCTTATACCTGATGGGCTGAACCTCTCAACCTGCAACACCTCCATGTAC 177
QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
DB 178 TTCTTCTTTTAACCTCTCTTTATATAGATTCTGTTCTTCCATCACTACCTACCTCCAAA 237
QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
DB 238 ATGCTGATGAGTTTGTCTCAAGAGAACATCATTTCTTCAAGGGGTATGACTCAG 297
QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120
DB 238 CTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 357
QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
DB 358 GACCGCTACGTCGACATCTGTAAACCATGTTGTATACAGACACATGATCTTCCAGGTG 417
QY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
DB 418 TGTTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 477
QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
DB 478 GGAAGCATATGAACTTCACTTCTGTCGACAACTGTCATCATCTTCACTTCTTCTGTCAG 537
QY 181 IleLeuProLeuLeuLeuLeuSerCysAsnGlySerTyrIleAsnValIleValIlePhe 200
DB 538 ATCTTCTCTCTCTGACCTCTCTCTGCAACGCTCTTCAACATGATGAGTGGTGTCTTT 597
QY 201 IleValIleThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
DB 598 ATTGGGTGGCTGTGACGTGGAATGCCATTTGTCACTGCTTATTTCTTATGCTTC 657
QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGlnGlyArgSerIleValIlePheSer 240
DB 658 ATCTCTCTCAACATCTTCAACATGATGTCATCAACATGATGTCATCAACATGATGTCAT 717
QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyValIlePheThrTyrLeu 260
DB 718 TGCAGTTCACATATATGATGATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 777
QY 261 LysProProSerIleLeuProLeuAspGlnGlyIleValSerSerLeuPheTyrThrThr 280
DB 778 AAACCCCTTTCATCTCTGCGCCCTGAGAGAGGAAAGTCTCTCTCTTCTTATACATA 837
QY 281 ValValPheMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValIleValAla 300
DB 838 ATAGTCCCGGATTAACCCATTAATCATATAGTCTTGGAACAAGATGTCAAAAGTTGCC 897
QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311
DB 898 CTGAGGAGAACTTTGGGAGAAAATCTTTCT 930

```

## RESULT 10

ABK37521  
ID ABK37521 standard; cDNA; 933 BP.

XX AC ABK37521;

XX DT 08-MAY-2002 (first entry)

XX DNA encoding G-coupled olfactory receptor #23.

KW Human; olfactory G-coupled receptor; sensory perception of odourant;  
KW odour composition; taste composition; gene; ss.

```

XX OS Homo sapiens.
XX XX
XX MN W0200198526-A2.
XX XX
XX PD 27-DEC-2001.
XX XX
XX PF 22-JUN-2001; 2001WO-US020122.
XX XX
XX PR 22-JUN-2000; 2000US-021812P.
XX PR 13-MAR-2001; 2001US-00804291.
XX XX
XX PA (SENO-) SENOMIX INC.
XX PI Zozulya S, Stryer L;
XX DR WPI: 2002-083330/11.
XX DR P-PSDB; AUB85162.
XX PT Representing sensory perception of one or more odourants for the
XX PT identification and design of tastes and odors comprises providing a
XX PT representative group of n olfactory receptors.
XX PS Example; Page 68; 182pp; English.
XX CC The invention relates to a method of representing sensory perception of
XX CC one or more odourants. The method comprises: (a) providing a
XX CC representative class of n olfactory receptors or ligand binding domains
XX CC (LBD) of these receptors; (b) measuring values X1 to Xn representative of
XX CC at least one activity of one or more odourants selected from: (i) binding
XX CC one or more odourants to the LBD of at least one of the n olfactory
XX CC receptors; (ii) activating at least one of the n olfactory receptors with
XX CC the one or more odourants; and (iii) blocking at least one of the n
XX CC olfactory receptors with the one or more odourants; and (c) generating a
XX CC representation of sensory perception from the values X1 to Xn. The
XX CC design and formulation of odour and taste compositions. ABK37499-ABK37754
XX CC and ABK37918-ABK37921 represent human olfactory G-coupled receptor coding
XX CC sequences and related PCR primers of the invention
XX SQ Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 6,09e-109 Length: 933
XX Score: 1273.50 Matches: 250
XX Percent Similarity: 87.46% Conservative: 22
XX Best Local Similarity: 80.35% Mismatches: 38
XX Query Match: 80.75% Indels: 1
XX DB: 6 Gaps: 1
XX
XX US-10-023-597-24 (1-311) x ABK37521 (1-933)
XX
QY 1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
DB 1 ATGGGAGCCCAAAAC--TCTTCTGTGACAGAGTTTACCTCGAAGGCTTAAACCCACAG 57
QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
DB 58 CCGGACACGCGGATCCCTCTTCTTCTGTTTCTGAGTTCTTCAACCGTCAACCGTGTG 117
QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
DB 118 GGAACCTGGGGCTTATACCTGATGGGCTGAACCTCTCAACCTGCAACACCTCCATGTAC 177
QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
DB 178 TTCTTCTTTTAACCTCTCTTTATATAGATTCTGTTCTTCCATCACTACCTACCTCCAAA 237
QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
DB 238 ATGCTGATGAGTTTGTCTCAAGAGAACATCATTTCTTCAAGGGGTATGACTCAG 297
QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120

```







Db	738	ATCCTTCCTCTCCTTAGCTCTCTCCGACACAGCTCTTACAGATAGCTGGTGGTCTT		797
Oy	201	IIeValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe		220
Db	798	ATTGTGGTGGCTTTACCGCTTGGATGGCCATTGCACTGCTTATTCTTATGGCCTC		857
Oy	221	IIleIleSerSerIleLeuLysValSerSerAlaGluGlyAspSerIleValAlaPheSerSer		240
Oy	241	CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu		260
Db	918	TGCAGTTCCACATATTTAGTTTCTCTTTCTTTGGTTCTGGTGGCTTCATGATATC		977
Oy	261	LysProProSerSerIleLeuProLeuAspGluGlyValSerSerLeuPheTyrThr		280
Db	978	AAACCCCTTTCATCTCGCCCTCGACGAGGAAAGTGCTCTCCGTTCTATACATCA		103
Oy	281	ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla		300
Db	1038	ATATGTCCTCCGTGTATACCACTTATCTATGCTTGAGAGACAGATGTCAAGTGGCC		109
Oy	301	LeuLysArgThrPheSerArgIleSerPheSer		311
Db	1098	CTGAGGAGAACTTTGGCGAGAAATAATCTTTCT		1130
RESULT 13				
ID	ABK97207	standard, cDNA; 976 BP.		
AC	ABK97207;			
DT	07-OCT-2002	(first entry)		
XX				
DE		Human G-protein coupled receptor (GPCR) GPCRX gene #1.		
XX				
KW		G-protein coupled receptor; receptor; GPCR; GPCRX; cardiomyopathy;		
KW		atherosclerosis; diabetes; cell signal processing; cancer; trauma;		
KW		metabolic pathway modulation; neuro-olfactory system; surgery;		
KW		neoplastic disorder; adenocarcinoma; lymphoma; prostate cancer;		
KW		uric acid cancer; immune response; acquired immunodeficiency syndrome; AIDS;		
KW		asthma; Crohn's disease; multiple sclerosis;		
XX		Albright hereditary osteodystrophy; gene therapy; gene; ss.		
OS		Homo sapiens.		
XX				
PN	WO200250117-A2.			
XX				
BD	27-JUN-2002.			
XX				
PF	18-DEC-2001; 2001WO-US049077.			
XX				
PR	18-DEC-2000; 2000US-0256635P.			
PR	21-DEC-2000; 2000US-0257876P.			
PR	04-JAN-2001; 2001US-0259743P.			
PR	10-JAN-2001; 2001US-0260718P.			
PR	12-JAN-2001; 2001US-0261498P.			
PR	24-JAN-2001; 2001US-0263689P.			
PR	08-FEB-2001; 2001US-0267464P.			
PR	22-FEB-2001; 2001US-0271212P.			
PR	14-MAR-2001; 2001US-0275946P.			
PR	23-MAR-2001; 2001US-0278150P.			
PR	18-APR-2001; 2001US-0284591P.			
PR	23-APR-2001; 2001US-0285718P.			
PR	19-JUN-2001; 2001US-0299327P.			
PR	16-AUG-2001; 2001US-0312902P.			
XX				
PA	(CURA-) CURAGEN CORP.			
XX				
PT	Padigaru M, Kekuda R, Colman SD, Spytek KA, Ballinger RA;			
PI	Vernet CAM, Li L, Shenoy S, Caeman SJ;			
XX				

DR WPI: 2002-528447/56.  
XX P-PSDB; ABG68134.

PT New G-protein coupled receptor polypeptides for treating or preventing  
PT cardiovascularity, atherosclerosis, diabetes, multiple sclerosis, acquired  
PT immunodeficiency syndrome or cancer in humans.

PS Claim 5; Page 79; 110pp; English.

CC The present invention relates to a new G-protein coupled receptor (GPCR).  
CC polypeptide. The GPCR polypeptide, GPCR nucleic acid and antibody are  
CC useful for treating, preventing or alleviating a GPCR-associated  
CC disorder or a pathological state in a subject, particularly a human. In  
CC particular, the disorder is cardiovascularity, atherosclerosis, diabetes, or  
CC a disorder related to cell signal processing and metabolic pathway  
CC modulation. The GPCR polypeptide and nucleic acid are also useful for  
CC diagnosing the presence of or predisposition to a disease associated with  
CC altered levels of GPCR, particularly cancer. The GPCR nucleic acid and  
CC polypeptide are especially useful in therapeutic or prophylactic  
CC applications for disorders of the neuro-endocrine system, e.g. those  
CC induced by trauma, surgery and/or neoplastic disorders. The DNA encoding  
CC the protein is useful in gene therapy for treating the above conditions.  
CC Furthermore, the nucleic acids and polypeptides are useful in treating  
CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune  
CC response, acquired immunodeficiency syndrome (AIDS), asthma, Crohn's  
CC disease, multiple sclerosis or Albright hereditary osteodystrophy. These  
CC are also useful in developing powerful assay system for functional  
CC analysis of various human disorders, as well as in diagnostic  
CC applications. The present nucleic acid sequence represents one of a  
CC collection (ABK97207-ABK97226) of human GPCR genes that encode the human  
CC GPCR proteins (AAU68134-AAU68153) of the invention

SQ XX Sequence 976 BP; 204 A; 259 C; 204 G; 309 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2,09e-108 Length: 976  
Score: 1268.00 Matches: 247  
Percent Similarity: 86.50% Conservative: 22  
Best Local Similarity: 79.42% Mismatches: 42  
Query Match: 80.41% Indels: 0  
DB: 6 Gaps: 0

US-10-023-597-24 (1-311) x ABK97207 (1-976)

OY 1 MetAlAlAGlunSensSerSeSeValThrgUnPheIleLeuAlaGlyLeuIIeHisGln 20  
Db : ::::: 8 ATGGCGACGCCAAAACCTCTGTGCAAGATTATCCCGAAGGCTTAACCCACAG 67  
OY 21 ProGLVLeuGIValProValPhePhePheLeuGlyPheTYrAlaValIThrValVal 40  
Db CGGGGACTGGGATCCCCCTCTTCTTCCTGTTCGGTTTCTACAGGTCACCGTG 127  
OY 41 GlyAsnLeuGlyLeuIIeLeuIIeGlyLeuAsnSerArgLeuHISilePrometTYr 60  
Db : ::::: 128 GGGAACCTGGCCTTGATAACCCTGTATGGGCTGAACCTCACCTGCACATCCCATGTAC 187  
OY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrTrpIIleLeuPolys 80  
Db TTCTTCTCTTTAACCTCTCTTAAATTTCTGTTTCTTCATTCATCACTACCTCCCAA 247  
OY 81 MetLeuMetSerPheValSerArgLysAsnIIeIleSerPheThrGlyCYSmetsSerGln 100  
Db ATGCTGATGATGTTTGCTCAAGAAGAACATCATTTCTTCACAGGATGATGATCTCAG 307  
OY 101 PhePhePhePheCysPhePheValAspSerGlySerPheIIeLeuSerAlaMetValGlu 120  
Db CTCTTCTTCTTCTGCTCTTCTTCTGCTCTGAGTCTTCATCCCTGACGATGGCTAT 367  
OY 121 AspArgTyValGlyIleCysAsnProLeuLeuTYrThrIIeThrMetSerProGlnVal 140  
Db GACCGCTACGTGGCCATCTGTAAACCATCTGTTGTACAGACACCAAGTCTTGCACAGTG 427  
OY 141 CysLeuLeuLeuLeuGlyValTyrlqyMetGlyValPheGlyAlaValAlaHisThr 160



```

Db 428 TGTTCCTCTCTTGTGTCCTATGGATGGGCTTGTCTGGGGCCATGGCCACACA 487
Qy 161 GYAANIIEVALPHELEUThrPheCYsAlaAspAsnLeuValaAsnHISYMetCYsASP 180
Db 488 GGAACACATATGACCTACCTCTCTGTCTGACAACTCTGCACTCATCTATCTATGTGAC 547
Qy 181 IIELEUPPLEULEUGLIEUSeCYsAsnGlySerTYrIleAsnValleuValIlePhe 200
Db 548 ATCTCTCTCTCTGACCTCTCCGCAACAGCTCTTACAGATGAGCTGTGTCTT 607
Qy 201 IIEVALIAlmValGlyIleGlyValProIleValIAlaValPheIleSerTYrGlyPhe 220
Db 608 ATTGTGCTGCTGTGACGTGAGATGCCATGTGCTACTGTCTTATTTCTATGCCCCC 667
Qy 221 IIELEUSERSerIleLeuArgValSerSerAlaGluGlyArgSerIlyAsaIlePheSer 240
Db 668 ATCTCTCTGACATCTACACACAGCTCTACAGAGGACAGCTCAAGCTTATGACT 727
Qy 241 CysSerSerTYrIleIleAlaValSerleuPheGlySerGlyAlaPheThrTYrLeu 260
Db 728 TGCAGTTCACATATGATGATTTCTTTCTTTGTTGTGTCTTGTCTTATGATCTC 787
Qy 261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerleuPheTYrThr 280
Db 788 AAACCCCTTCCATCTCTCCCTCGAGCAGAGGAAAGTCTCTCTCTATACATA 847
Qy 281 ValValProMetPheAsnProLeuIleTYrSerleuArgAsnLysAspValLysLeuAla 300
Db 848 ATAGTCCCGCGTAAACCATTAATCTATGCTTGAGAGAACAGATGTCMAAGTGTCC 907
Qy 301 LeuLysArgThrPheSerArgIleSerPheSer 311
Db 908 CTGAGGAGAACTTTGGGCAAGAAATCTTTTCT 940

```

RESULT 14  
AAD60328  
ID AAD60328 standard; DNA; 976 BP.

AC AAD60328;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human G-protein coupled receptor (GPCR) GPCR DNA #1.  
XX  
KW Human G-protein coupled receptor; GPCR; stroke; obesity; virucide;  
KW multiple sclerosis; Alzheimer's disease; graft-versus-host disease;  
KW endometriosis; tissue typing; gene therapy; vaccine; antibacterial;  
KW retinal disease; infectious disease; bulimia; receptor; fungicide;  
KW Parkinson's disease; hypertension; acute heart failure; infection;  
KW cardiomyopathy; atherosclerosis; diabetes; detectability disorder;  
KW cancer; signal transduction pathway disorder; metabolic disorder;  
KW developmental disorder; osteoporosis; protozoasidae; asthma; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 8..943  
FT /\*tag= a  
FT /product= "Human GPCR protein"

XX US2003100491-A1.  
XX  
PD 29-MAY-2003.  
XX  
PF 18-DEC-2001; 2001US-00024399.  
XX  
PR 18-DEC-2000; 2000US-0256635P.  
PR 21-DEC-2000; 2000US-0257976P.  
PR 04-JAN-2001; 2001US-025743P.  
PR 10-JAN-2001; 2001US-0260718P.  
PR 12-JAN-2001; 2001US-0261498P.  
PR 24-JAN-2001; 2001US-0263689P.

```

PR 08-FEB-2001; 2001US-0267464P.  
PR 22-FEB-2001; 2001US-0271021P.  
PR 14-MAR-2001; 2001US-0275946P.  
PR 23-MAR-2001; 2001US-0278150P.  
PR 18-APR-2001; 2001US-0284591P.  
PR 23-APR-2001; 2001US-0285718P.  
PR 19-JUN-2001; 2001US-0295327P.  
PR 16-AUG-2001; 2001US-0312902P.  
XX  
PA (PADI/) PADIGARU M.  
PA (KEKU/) KEKUDA R.  
PA (COLM/) COLMAN S. D.  
PA (SPRYT/) SPRYTER K. A.  
PA (BALU/) BALLINGER R. A.  
PA (VERN/) VERNET C. A. M.  
PA (LILL/) LI L.  
PA (SHEN/) SHENOY S. G.  
PA (CASW/) CASMAN S. J.  
PI Padigaru M., Kekuda R., Colman S.D., Spryter K.A., Ballinger R.A.  
PI Vernet C.A.M., Li L., Shenoy S.G., Casman S.J.  
XX  
DR MPI; 2003-755210/71.  
XX P-PSDB; AAE39659.  
XX  
PT New G-protein coupled receptor (GPCR) polypeptides and nucleic acids,  
PT useful for diagnosing, preventing or treating GPCR-associated disorders,  
PT e.g. cardiomyopathy, atherosclerosis, cancer or diabetes, and in  
PT pharmacogenomics.  
XX  
PS Claim 8; Page 32; 46pp; English.  
XX  
CC The invention relates to G-protein coupled receptor (GPCR) polypeptides  
CC (designated as GPCR) and nucleic acid sequences. GPCR polypeptides,  
CC nucleic acids and antibodies are useful in diagnosing, preventing or  
CC treating GPCR-associated disorders, such as cardiomyopathy, diabetes,  
CC atherosclerosis and cancer. These may also be used for treating or  
CC preventing other diseases like developmental disorders, taste and scent  
CC detectability disorders, signal transduction pathway disorders, retinal  
CC diseases, metabolic disorders (e.g. obesity), infectious diseases (e.g.  
CC bacterial, fungal, protozoal or viral infections), bulimia, asthma,  
CC Parkinson's disease, hypertension, acute heart failure, osteoporosis,  
CC multiple sclerosis, Alzheimer's disease, stroke, graft-versus-host  
CC disease or endometriosis. GPCR nucleic acid is used as hybridisation  
CC probes, in chromosome mapping, tissue typing, preventive medicine and  
CC pharmacogenomics. It is also useful in gene therapy. GPCR is useful as  
CC vaccines. The present sequence is human GPCR DNA  
XX  
SQ Sequence 976 BP; 204 A; 259 C; 204 G; 309 T; 0 U; 0 Other;  
XX  
Alignment Scores:  
Pred. No.: 2,096-108 Length: 976  
Score: 1268.00 Matches: 247  
Percent Similarity: 86.50% Conservative: 22  
Best Local Similarity: 79.42% Mismatches: 42  
Query Match: 80.41% Indels: 0  
DB: 9 Gaps: 0  
US-10-023-597-24 (1-311) x AAD60328 (1-976)  
Qy 1 MetaAlaIAGLUAISerSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20  
Db 8 ATGGGCCAGCCCAAAACCTTCTGTGACAGATTAATCCGAGAGCTTAACCCACAG 67  
Qy 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTYrAlaValThrValAl 40  
Db 68 CCGGACCTGGCGAGATCCCTCTCTCTCTGTTCTGTGGGTTCTACAGGTCACCGTGTG 127  
Qy 41 GYAANIIEUGLIEUleuIleleuIleleuIleGlyLeuAsnSerArgLeuHisIleProMetTYr 60  
Db 128 GGAACCTGTGATTAACCTGATGGGCTGTAACCTGACACCTGACACCTCCATGTAC 187  
Qy 61 PhePheProPheAsnLeuSerleuValAspPheSerThrThrIleIleProLys 80

```



```
Db      16 ATGGCAGCCAAAAC---TCTTCTGTGACAGATTATCTCGAAGCTTAAACCACAG 72
Qy      21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
Db      73 CCGGGACTCGGAGTCCCTCTTCTTCTGCTTCTGAGTTTCTACACGGTCACTGGTGG 132
Qy      41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
Db      133 GGGAACTGGGCTGTGTAACCTGATGGGCTGAATGCTCACTGACACTCCCATGTATC 192
Qy      61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
Db      193 TTTCTTCTTTTAACTCTCTTAAATAGATTCTGTCTCTCCATCCTACATCCTCCCAA 252
Qy      81 MetLeuMetSerPheValSerArgIleAsnIleIleSerPheThrGlyCysMetSerGln 100
Db      253 ATGCTGATGATGATTGTGCTCAAGAGAAACATCATTTCTTCACAGSGGTATGACTCAG 312
Qy      101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120
Db      313 CTCTTCTTCTTCTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
Qy      121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
Db      373 GACCCCTAGTGGCCATCTGTACCCACTGTTGTACACAGTCACCATGTCTTGGCCAGTG 432
Qy      141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
Db      433 TGTTCCTCTCTTCTTGTGGTGCTGATGGGATGGGCTTGTGGGGCCATGGCCACACA 492
Qy      161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
Db      493 GGAACCATATGAACTTACCTCTGTGTGACCACTTGTCAATCATTCATTCATGTGTGAC 552
Qy      181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
Db      553 ATCCTTCTCTCTCTGAGCTCTCTGCAACAGCTCTTACATGAATGAGCCGCTGCTTT 612
Qy      201 IleValIleThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
Db      613 ATGTGGTGGCTGTGAGCTTGAATGCCATGTGCACTGTCTTATTTCTATGCCCTC 672
Qy      221 IleLeuSerSerIleLeuArgValSerSerAlaGlyIleArgSerIleValAspSerSer 240
Db      673 ATCCTCTCCAGATTCTACACACAGATTCTACAGAGGCAAGTCCAAAGCTTTAGTACT 732
Qy      241 CysSerSerTyrIleIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
Db      733 TGCAGTTCCACATATATGATTCTCTTCTTCTTCTGTTCTGTTCTGTTCTCATGTATCTC 792
Qy      261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheIleThrThr 280
Db      793 AAACCCCTTCCATCTCTGCTCCCTCGAGCAGGAAAGTCTCTCTCTGTTCTATACCATA 852
Qy      281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
Db      853 ATAGTCCCGGTGTAAACCATTAATCATGTGAGGAAACAAGATGTCAAAGTTGCC 912
Qy      301 LeuLysArgThrPheSerArgIleSerPheSer 311
Db      913 CTGAGAGAACTTTGGGCAAAAAATCTTTTCT 945
```

Search completed: September 30, 2004, 10:09:15  
Job time : 455 secs



Db 425 AATGCTGATGATTTTGTCTCAAGAGAAACATTATTTCTTCACAGGTGTATGAGTCA 366  
 QY 100 nphphphepheCysphepheValpheserGluSerPheIleuSerAlaMetValGI 120  
 Db 365 GTTCTCTCTCTTCTGTTCTTCTGTTCTTCTGATCTCTTAT-CTGCGCGCATGGCGTA 307  
 QY 120 uAspArgTyValGlyIleCysAsnProLeuLeuThrIleTherMetSerProGlnVal 140  
 Db 306 GGACCGGTAGCTGGCATCTGTATCCATTTGTGTACACGTCACATGCTCCCGCAGGT 247  
 QY 140 lCysleuLeuLeuLeuGlyValIlyrGlyMetGlyValPheGlyAlaAlaIleHstH 160  
 Db 246 GTGTTGTCTCTTCTTACTGGGTGTCTACGGGATGGGGTTTGGGGCTGGGCTCATAC 187  
 QY 160 rGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHstIlyrMetCysAs 180  
 Db 186 AGGAATATATGTTTCTTCTACCTTTTGTGACAGACACTGTGATCATACATGATGTGA 127  
 QY 180 pIleuProLeuLeuGluLeuSerCysAsnGlySerTyIleAsnValIleuValIleph 200  
 Db 126 CATCTTCCCTCTTCTTGTAGCTCTCTCTGCAAGGCTCTTACATAATGCTCTGATCATCTT 67  
 QY 200 eIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyGlyPh 220  
 Db 66 TATGTTGTGACCGTTGGCATTTGGGTG-CCCATTTGTGCGCTTTTATCTTATGTTT 8  
 QY 220 eIleLeu 222  
 Db 7 TATCTCT 1  
 RESULT 2  
 US-09-016-434-1114  
 ; Sequence 1114, Application US/09016434  
 ; Patent No. 6500938  
 ; GENERAL INFORMATION:  
 APPLICANT: Janice Au-Young  
 APPLICANT: Jeffrey J. Seilhamer  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 NUMBER OF SEQUENCES: 1490  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/016.434  
 FILING DATE: HEREMITH  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0002 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1114:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 945 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: G1336040  
 US-09-016-434-1114  
 Alignment Scores:  
 Pred. No.: 2,936-67  
 Score: 705.50  
 Percent Similarity: 65.59%  
 Best Local Similarity: 44.69%  
 Query Match: 44.74%  
 DB: 4  
 Gaps: 3  
 US-10-023-597-24 (1-311) x US-09-016-434-1114 (1-945)  
 QY 5 AenSerSerValThrGluPheIleuAlaGlyLeuIleHstGlnProGlyLeuGln 24  
 Db 19 AACCTACAGTGGTCTGACGATGATTTCTATAGGTTTCCAACTCCGCTGAGCTGAG 78  
 QY 25 ValProValPhePheLeuPheLeuGlyPheTyAlaValThrValValGlyAsnLeuGly 44  
 Db 79 ATGTGCTGTTCTGATGTTTCTGACATTGATGATTAATTCGATAGGAGACATTGGA 138  
 QY 45 LeuIleIleLeuIleGlyLeuAsnSerArgLeuHstIlePrometTyPhePheProPhe 64  
 Db 139 TTGATGCTGTTGATGAGATTTGATTCACCTTCAACCCCATGATATTTTCTTACG 198  
 QY 65 AenLeuSerLeuValAspPheSerPheSerThrThrIleIleProIleMetLeuMetSer 84  
 Db 199 AACCTATCATTTGTAGACCTTGTGCTTCTGACATGTTTCCCAAAATGCTGTCAT 258  
 QY 85 PheValSerArgGlyAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhe 104  
 Db 259 TTCTCTCGGAGAACAAATCTATTTCTTATATGAGTGTGCGCTTATTTTCTT 318  
 QY 105 CysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyVal 124  
 Db 319 TGTACTTTGGACAGATACAAATCTCTCATCTGCGCCGACATGATGATGCTATGTC 378  
 QY 125 GlyIleCysAsnProLeuLeuThrIleTherMetSerProGlnValCysLeuLeuLeu 144  
 Db 379 GCCATCTGTAACCTTTATTTGATGACAGTGTGATGTTGAGGGCATGTATGCGGTG 438  
 QY 145 LeuLeuGlyValIlyrGlyMetGlyValPheGlyAlaValAlaHstIleGlyAsnIleVal 164  
 Db 439 ATGTGCTGTATACCTTGAAGGCAACATGATGCTCTGCTGACACATCTTGGCTT 498  
 QY 165 PheLeuThrPheCysAlaAspAsnLeuValAsnHstIlyrMetCysAspIleLeuProLeu 184  
 Db 499 ATTCGAATATTTGACAAATAATGTTATATCATTTTCTGTGACCTCTCCCTG 558  
 QY 185 LeuGluLeuSerCysAsnGlySerTyIleAsnValIleuValIlePheIleValIleThr 204  
 Db 559 CTTAACTATCTCTGACACGACACACAACTAATATG-----TGGCTCTCCACCA 609  
 QY 205 ValGlyIleGlyValProIleValAla-----ValPheIleSerTyGlyPheIle 221  
 Db 610 TACGGCAGCTCAGTGAATCATTTGTTTATCATCATCATCATCTCTTACTTTTTCATT 669  
 QY 222 LeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerTyAlaPheSerSerCys 241  
 Db 670 CTTCTCTCAGTTTAAAGATCCCTCTTCTGAGTGAAGAAAGACCTTTTCTACATGC 729  
 QY 242 SerSerTyIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyLeuLys 261  
 Db 730 GCCTCTACACTGACTGACGTGACATCAACAGGAGATCTCTCTTATTTACTCAGG 789  
 QY 262 ProProSerIleLeuProLeuAspGlnIlyrValSerLeuLeuPheTyThrVal 281  
 Db 790 CCACACTACCTGATATTTCCAAACATGATTAATATCTAGAGTTTACACATTTTC 849  
 QY 282 ValProMetPheAsnProLeuIleTySerLeuArgLeuAspValIlyrLeuAlaLeu 301

```

Dh      850  ATTCAGTGTGATACCGTTGATTATTAAGTTGGAAGAAATAAGATTAAGATGACGCT
Oy      302  LyeArGThrPhe---SerArgLLeSerPheSer 311
Db      910  GAGAAAGTTCTTAAGATCAAGGATGAGATCTTCA 942

RESULT 3
US-09-016-434-1413
: Sequence 1413 Application US/09016434
: Patent No. 6500938
: GENERAL INFORMATION:
: APPLICANT: Janice Au-Young
: APPLICANT: Jeffrey J. Sellhammer
: TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
: TITLE OF INVENTION: PATHWAY GENE EXPRESSION
: NUMBER OF SEQUENCES: 1490
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
: STREET: 3174 PORTER DRIVE
: CITY: PALO ALTO
: STATE: CALIFORNIA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/016,434
: FILING DATE: HEREWITH
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Zeller, Karen J.
: REGISTRATION NUMBER: 37,071
: REFERENCE/DOCKET NUMBER: PA-0002 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (650) 855-0555
: TELEFAX: (650) 845-4166
: INFORMATION FOR SEQ ID NO: 1413:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1282 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GENBANK
: CLONE: g516319
: US-09-016-434-1413

Alignment Scores:
Pred. No.: 6,85e-59 Length: 1282
Score: 630.00 Matches: 122
Percent Similarity: 60.87% Conservative: 60
Best Local Similarity: 40.80% Mismatches: 117
Query Match: 39.95% Indels: 0
DB: Gaps: 0

US-10-023-597-24 (1-311) x US-09-016-434-1413 (1-1282)

Oy      5  AsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGln 24
Db      180  AATGACACAGCTCATTTGCTGAGTTCATCTGCTGGGGCTTGCTGAGGCCCGACGGCTGCAG 235
Oy      25  ValProValPhePheLeuPheLeuGlyPheTyrAlaValAlaThrValValGlyAsnLeuGly 44
Db      240  CCAAGTTGTCCTTTGGCTCTTCCCTCTTGCGCTACCTGCTACGCTACGAGGCGACACTCAGC 295
Oy      45  LeuIleLeuLeuGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPhe 64

```

PRIOR FILING DATE: 2000-03-31  
 NUMBER OF SEQ ID NOS: 13  
 SOFTWARE: Pct\_Files Version 2.0  
 SEQ ID NO 10  
 LENGTH: 1062  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: (1)..(1062)  
 US-09-668-680-10

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
4.95e-58	621.00	1062	128
Percent Similarity:	61.36%	Conservative:	61
Best Local Similarity:	41.56%	Mismatches:	117
Query Match:	39.38%	Indels:	2
DB:	4	Gaps:	2

US-10-023-597-24 (1-311) x US-09-668-680-10 (1-1062)

```

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleuAlaGlyLeuIleHISGln 20
DB 124 ATGGAAGCAGGAAACCAACAGATTTTATGATTTATCTTCTCGAGCTCTGAGGAT 183
QY 21 ProGlyLeuGlnValProValPhePheLeuGlyPheTyrAlaValThrVal 40
DB 184 CCAGACTACAGCCGCTTCAATTTGGGCTGTCCTCCATGATGACCTGGGAGCGTCTG 243
QY 41 GlyAlaLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
DB 244 GGAACACCTGCTCATCTCTGCGCATCTGCTGATCTCCACCTCCACACCCCATCTAC 303
QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProGly 80
DB 304 TTCTTCTCTCCACCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
QY 81 MetLeuMetSerPheValSerArgGlyAsnIleIleSerPheThrGlyCysMetSerGln 100
DB 364 ATGCTGGTGAACATCCACACCGAGAACAAAGCCATCTCTCAATGATGATGATGATGATG 423
QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
DB 424 GTCTATTTCTCCATCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 483
QY 121 AspArgTyrValIleGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
DB 484 GACCGGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaAlaHisThr 160
DB 544 TGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 603
QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
DB 604 TCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 663
QY 181 IleuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIleValIlePhe 200
DB 664 CTGACGTACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
QY 201 IleValValThrValGlyIleGlyVal---ProIleValAlaValPheIleSerTyrGly 219
DB 724 TTTATGAGGGGTG---CTGGGGGTTTTCCTCTTGGGATTTTCTCTTATTTCA 780
QY 220 PheIleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerTyrAlaPheSer 239
DB 781 CGATTTGCTTATTCATTAAGAAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 240 SerCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyr 259
DB 841 ACCTGCTGCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
  
```

```

QY 260 LeuLysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThr 279
DB 901 TTCATTTCTGCGTGAATCACTCTTCCGAAACAAATCTCCGGGCTCGGTGATGACACT 960
QY 280 ThrValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeu 299
DB 961 GTGGTACCCCATGTTGAACCTTCTTCACTTCACTGAGGAGAACGATGTGAAGGCA 1020
QY 300 AlaLeuLysArgThrPheSerArg 307
DB 1021 GCCCTGGGAGTCTCTCTGAGCAGG 1044
  
```

## RESULT 5

US-09-016-434-1312  
 Sequence 1312, Application US/09016434  
 Patent No. 6500938

## GENERAL INFORMATION:

APPLICANT: Janice Au-Young  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 NUMBER OF SEQUENCES: 1490  
 CORRESPONDENCE ADDRESS:

ADDRESS: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO

STATE: CALIFORNIA  
 COUNTRY: USA

ZIP: 94304

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for windows/MS-DOS 6.2

CURRENT APPLICATION DATA: US/09/016,434

## CLASSIFICATION:

FILING DATE: HEREWITH

## PRIORITY INFORMATION:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0002 US

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1312:

SEQUENCE CHARACTERISTICS:

LENGTH: 1854 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: 932085

US-09-016-434-1312

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
1.8e-57	619.00	1854	131
Percent Similarity:	62.36%	Conservative:	58
Best Local Similarity:	43.23%	Mismatches:	114
Query Match:	39.25%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x US-09-016-434-1312 (1-1854)

```

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleuAlaGlyLeuIleHISGln 20
DB 124 ATGGAAGCAGGAAACCAACAGATTTTATGATTTATCTTCTCGAGCTCTGAGGAT 183
  
```



Db 320 ATGATGGGACAAATCAACACGACATCTGACATTCCTGCTGAGGCTGCCATCCAA 379  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValAl 40  
 Db 380 CCAGAGCAGCAAAACCTGTGCTATGCGCTTTCGGCCATGATCTTACCAACCTCTCG 439  
 QY 41 GlyAsnLeuGlyLeuLeileuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60  
 Db 440 GGGAACTCTCATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 499  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerPheThrThrLeuLeu 80  
 Db 500 TTGTTCTCAGCAACTGCTCTCTCTGACCTCTGCTCTCTCTCTCTCTCTCTCTCT 559  
 QY 81 MetLeuMetSerPheValSerArgLysAsnLeileuSerPheThrGlyCysMetSerGln 100  
 Db 560 TTGTTACAGAACATGAGAACCAAGACCATTCATCCCTATGCGGACTGCTGACCCA 619  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheLeuLeuSerAlaMetValGlu 120  
 Db 620 ATGACTTCTCTCTGTTATTTGGAGACCTGAGAGCTTCTCTCTGCTGCTGCTGCT 679  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThmMetSerProGlnAla 140  
 Db 680 GACCGCTATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 739  
 QY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 Db 740 TGTCTGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisThrCysAsp 180  
 Db 800 TTACTCATGGCCAGGTTGTTGTTTGTGAGACCAATGATCCCACTTTTCGTGAT 859  
 QY 181 IleLeuProLeuLeuGlnLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 Db 860 ATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 919  
 QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 Db 920 ATCATGGAGGCGCTCATCTTTCATCATCCATCTCTCATCTCTGCTGCTGCTGCT 979  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGlyIleGlyValSerIleAspPheSer 240  
 Db 980 ATGTCTCTCTCATCTCTCAAGGCTCTCTCTCAAGGCTCTCTCTCTCTCTCTCT 1039  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260  
 Db 1040 TGTGCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1099  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheThrThr 280  
 Db 1100 TGTCTCATGACTAATGTTCTACTTAAGGACATGTCATGCTATGATGATGACCTG 1159  
 QY 281 ValValProMetPheAsnProLeuLeuIleTyrSerLeuArgAsnLysValIleLeuAla 300  
 Db 1160 GTGACCCCATGCTGTAACCCCTTCTCATCTACAGCTGAGAGACAGACATGAGGAGGC 1219  
 QY 301 LeuLysArg 303  
 Db 1220 CTGAGCAGA 1228  
 RESULT 6  
 US-08-748-506-6  
 ; Sequence 6, Application US/08748506  
 ; Patent No. 6159707  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ronnett et al.  
 ; TITLE OF INVENTION: NOVEL SPERM RECEPTORS  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Leydig, Voigt & Mayer, Ltd.  
 ; STREET: Two Prudential Plaza, Suite 4900

CITY: Chicago  
 STATE: IL  
 COUNTRY: US  
 ZIP: 60601-6780  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/748,506  
 FILING DATE: 08-NOV-1996  
 CLASSIFICATION: 435  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 60/033,751  
 FILING DATE: 09-NOV-1995  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION: 74940  
 TELEPHONE: 312-616-5700  
 TELEFAX: 312-616-5700  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 966 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 US-08-748-506-6  
 Alignment Scores:  
 Pred. No.: 1,51e-57 Length: 966  
 Score: 616.00 Matches: 130  
 Percent Similarity: 59.47% Conservative: 49  
 Best Local Similarity: 43.19% Mismatches: 122  
 Query Match: 39.06% Indels: 0  
 Gaps: 0  
 US-10-023-597-24 (1-311) x US-08-748-506-6 (1-966)  
 QY 4 GluAsnSerSerSerValThrGluPheIleuAlaGlyLeuIleHisGlnProGlyLeu 23  
 Db 28 GAGAAATGTTGCTGTCAACGCTTTCATTTGCCAAGTCTCTGAGCCCTCGAGAA 87  
 QY 24 GlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValAlaGlyAsnLeu 43  
 Db 88 TGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 147  
 QY 44 GlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePhePro 63  
 Db 148 CTCATAGCCCTTGTGTTGTTGATCCAGTCCATCTTACACACCCCAATGATCTTCTCTG 207  
 QY 64 PheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMet 83  
 Db 208 GCCAATGTTGCTCTCGAGATGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 267  
 QY 84 SerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhe 103  
 Db 268 AGCTTGTGAGTGGGCCCAAGAGATCTCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 327  
 QY 104 PheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGluAspArgTyr 123  
 Db 328 TTCACTTTTGGCANAAGTGAAGTGTGCTATGTCAGACCAATGCTTGAACCGCTGC 387  
 QY 124 ValGlyIleCysAsnProLeuLeuTyrThrIleThmMetSerProGlnValCysLeuLeu 143  
 Db 388 ATGGCATATGCTCCCACTCCATGACCAACCGAATGAGTGTGAGTGTGAGTGTGCCAT 447  
 QY 144 LeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIle 163  
 Db 448 TTGGCATTTGTTTCAATGGGAAATGGAGATGATAGTATGATGCTGGAGACAGCAATTTTAT 507

```

QY 164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuPro 183
DB 508 TNCCTCTGAACCTTGTGGACCTTGTGAGATAGACCACTTCTGTGGACCTTCCACT 567
QY 184 LeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIlePheIleValI 203
DB 568 CTCCTGGACCTTGCCTGTGGATATCATCCAAAACGAGGCTGCATCTTTGTGGACGA 627
QY 204 ThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPheIleLeuSer 223
DB 628 ATCCCTGTATATCATGATGCAATTTTGGTGAATCCTTTATCTTATGTCAGAAATCTCGTT 687
QY 224 SerIleLeuArgValSerSerIleGluGlyArgSerIleValAlaPheSerSerCysSerSer 243
DB 688 GCAGTCTGGTGATACCTTCACTGAGGGGCGCCATAAGCTCTTCCACCTGTTCTCA 747
QY 244 TyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrIleLeuPro 263
DB 748 CACCTACTGTGATGACACTCTTTTATGCTGTGCTCTTACCTATTGAGGCCAAG 807
QY 264 SerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThrValValPro 283
DB 808 TCTAGCCACTACCGAGATGACAACTTTGGCCCTTCTACACAGCAGTACATCC 867
QY 284 MetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArg 303
DB 868 ATGCTGAACCTATCATCTATACAGTCTTAAGAAACAAGATCAAGGACGACTAGAGA 927
QY 304 Thr 304
DB 928 ACT 930

```

## RESULT 7

US-08-748-506-5  
Sequence 5, Application US/08748506

```

PATENT INFORMATION:
GENERAL INFORMATION:
APPLICANT: Romelt et al.
TITLE OF INVENTION: NOVEL SPERM RECEPTORS
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leydig, Voigt & Mayer, Ltd.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: IL
COUNTRY: US
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/748,506
FILING DATE: 08-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,751
FILING DATE: 09-NOV-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 74940
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5600
TELEFAX: 312-616-5700
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-748-506-5

```

Alignment Scores:  
Pred. No.: 2,48e-57  
Score: 614.00  
Percent Similarity: 59.47%  
Best Local Similarity: 42.86%  
Query Match: 38.93%  
DB: 3  
Gaps: 0

US-10-023-597-24 (1-311) x US-08-748-506-5 (1-966)

```

QY 4 GluAsnSerSerValThrGluPheIleValIleGlyLeuIleGlnProGlyLeu 23
DB 28 GAGATAGTGTACCTGTCAACACTTTGGATTTGGCAAGTCTCTGAGGCTCCGAGAA 87
QY 24 GlnValProValPhePheLeuLeuGlyPheTyrAlaValThrValGlyAsnLeu 43
DB 88 TGCCTCTCTCTTCAACCTCATCCTTTCATGTTCTTAGATCATCAAGGAATCT 147
QY 44 GlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePhePro 63
DB 148 CTCATAGTCTCTGCTATTGTTGACAGTCCATCTCTACACACCCCATGACTTCTTCTG 207
QY 64 PheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuVal 83
DB 208 GCCAAGTGTCTCTCTGAGAGATGCGATACCTTCTGTGATACCAAGATGCTGCAG 267
QY 84 SerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhe 103
DB 268 AGCCTTGATGAGAGCCAGACAGATCTCTGGAGGAGATGCCACACAGATGTTT 327
QY 104 PheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGluAspArgTyr 123
DB 328 TTGGATTTTGGTATATGATGATGATGCTGATGAGCATGATGAGCATGAGCCTTGGACG 387
QY 124 ValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnValCysLeuLeu 143
DB 388 ATGGCCATATGCTCCCACTCCATATGCAACCCGAAATGATCGAGATATGCCAT 447
QY 144 LeuLeuGlyValIleTyrGlyMetGlyValPheGlyAlaValAlaHisThrGlyAsnIle 163
DB 448 TTGGCATATGTTTCAATGGAGATGGATGATGATAGTATGATCGGAGCAACAATTTT 507
QY 164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuPro 183
DB 508 TCTCTCTGAACCTTGTGGACCTTGTGAGATAGACCACTTCTGTGGACCTTCCACT 567
QY 184 LeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIlePheIleValIleValI 203
DB 568 CTCCTGGACCTTGCCTGTGGATATCATCCAAAACGAGGCTGCATCTTTGTGGACGA 627
QY 204 ThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPheIleLeuSer 223
DB 628 GTCTCTGATATCATGATGCAATTTTGGTGAATCCTTTATCTTATGTCAGAAATCTCGTT 687
QY 224 SerIleLeuArgValSerSerIleGluGlyArgSerIleValAlaPheSerSerCysSerSer 243
DB 688 GCAGTCTGGTGATACCTTCACTGAGGGGCGCCATAAGCTCTTCCACCTGTTCTCA 747
QY 244 TyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrIleLeuPro 263
DB 748 CACCTACTGTGATGACACTCTTTTATGCTGTGCTCTTACCTATTGAGGCCAAG 807
QY 264 SerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThrValValPro 283
DB 808 TCTAGCCACTACCGAGATGACAAATCTTGGCCCTTCTACACAGTACATCATCC 867
QY 284 MetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArg 303
DB 868 ATGCTGAACCTATCATCTATACAGTCTTAAGAAACAAGATCAAGGACGACTAGAGA 927
QY 304 Thr 304

```

Db 928 ACT 930

RESULT 8  
US-08-748-506-8  
Sequence 8, Application US/08748506  
Patent No. 6159707

GENERAL INFORMATION:  
APPLICANT: Rommett et al.  
TITLE OF INVENTION: NOVEL SPERM RECEPTORS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leydig, Voigt & Mayer, Ltd.  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: IL  
COUNTRY: US  
ZIP: 60601-6780

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/748,506  
FILING DATE: 08-NOV-1996  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 60/033,751  
FILING DATE: 09-NOV-1995

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
REFERENCE/DOCKET NUMBER: 74940  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-616-5600  
TELEFAX: 312-616-5700

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 966 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-748-506-8

Alignment Scores:  
Pred. No.: 1,11e-56 Length: 966  
Score: 608.00 Matches: 130  
Percent Similarity: 59.33% Conservative: 48  
Best Local Similarity: 43.33% Mismatches: 132  
Query Match: 38.55% Indels: 0  
Gaps: 0

US-10-023-597-24 (1-311) x US-08-748-506-8 (1-966)

QY 4 GUAASerSerserValthrGupheileuAlaGlyLeuileHisGlnProGlyLeu 23  
DB 28 GAGAAAGATTGCTGTCTCAACGCTTTTGCATTGCCAAGCTTCTGAGGCTCCCTGAGANA 87  
QY 24 GlnValProValPhePheLeuPheLeuGlyPheTYrAlaValThrValValGlyAsnLeu 43  
DB 88 TGCTTCTCTGCTTCACTCATCTTCTCATGTTCTTTGATCTACACAGAGAAATGCT 147  
QY 44 GlyLeuileileleuileGlyLeuAsnSerArgLeuHisIleProMetTYrPhePhePro 63  
DB 148 CTCATAGCCCTTGCTATTGTATCCAGTCCATCTACACACCCCATGACTTCTTTCTG 207  
QY 64 PheAsnLeuSerleuValAspPheSerPheSerThrThrIleIleProGlyMetLeuMet 83  
DB 208 GCCAACTTGTCTCTCTGAGATGGCTACTTGTCTGTATACCCCAAGATGCTGCAG 267  
QY 84 SerPheValSerArgIysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhe 103  
DB 268 AGTCTGTAGTGAAGCCGAGAGATCTTTCAGTGGAGATGTCACACACAGATGTTTTC 327

QY 104 PheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlyAspArgTYr 123  
DB 328 TTCATATTCTTGTATATGAGATGCTGCTCATTTGACACCAAGCCTTTGACCGCTAT 387  
QY 124 ValGlyIleCysAsnProLeuLeuTYrThrIleMetSerProGlnValCysLeuLeu 143  
DB 388 ATGGCTATATGTTCCCACTCCATGACACCCCAAGATGCTGAGATGAGGCCAC 447  
QY 144 LeuLeuGlyValTYrGlyMetGlyValPheGlyAlaValAlaIleThrGlyAsnIle 163  
DB 448 TTGGCAATTGTTATGATGGATGATGATGATGATGATGATGATGATGATGATGAT 507  
QY 164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTYrMetCysAspIleLeuPro 183  
DB 508 TTCCTCTGAACCTTCTGACCCCTGTGATATGACCACTTCTGTGATCTTCCACT 567  
QY 184 LeuLeuGlyLeuSerCysAsnGlySerTYrIleAsnValLeuValIlePheIleVal 203  
DB 568 CTCCTGGCACTTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGAT 627  
QY 204 ThrValGlyIleGlyValProIleValAlaValPheIleSerTYrGlyPheIleLeuSer 223  
DB 628 GTCTCTGCAATATCAACCCCTTTTGTGATCATCTTATCTTATGACAAATCTCTG 687  
QY 224 SerIleLeuArgValSerSerAlaGlyArgSerIleValPheSerSerCysSerSer 243  
DB 688 GCAGTGTCTGTGATGCTTCACTGAGGGGCGCACAAAGCCTTTTCACTGTTCTCTC 747  
QY 244 TYrIleIleAlaValSerleuPhePheGlySerGlyAlaPheThrTYrLeuIlePro 263  
DB 748 CACTACTTGTAGTACACACTTTTATGCTCAGAGATCTGTACCTTTGAGGCCCTAAG 807  
QY 264 SerIleLeuProLeuAspGlnGlyValSerSerleuPheTYrThrValValPro 283  
DB 808 TCTAGCCACTCACAGGAATGACAACTTGTGCTCTTCTTACACAGGAGCATCC 867  
QY 284 MetPheAsnProLeuIleTYrSerleuArgAsnIleValIleValIleValIleVal 303  
DB 868 ATGTTGAACCTTATCATCTATGTTTATGAGAAACAAGATGTCACAGCAGCATGAGAGA 927

RESULT 9  
US-09-546-986A-5  
Sequence 5, Application US/09546986A  
Patent No. 6635741

GENERAL INFORMATION:  
APPLICANT: Powers, Scott  
APPLICANT: Yang, Jianxin  
APPLICANT: Cutler, Gene  
APPLICANT: Tulaxik Inc.  
TITLE OF INVENTION: No. 6635741e1 G-Protein Coupled Receptors  
FILE REFERENCE: 018781-004720US  
CURRENT APPLICATION NUMBER: US/09/546,986A  
FILE FILING DATE: 2002-04-30  
PRIORITY FILING DATE: 2000-03-14  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 5  
LENGTH: 1351  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (86)..(1108)  
OTHER INFORMATION: human breast cancer amplified G-protein coupled  
OTHER INFORMATION: receptor 3 (BCA-GPCR-3)  
US-09-546-986A-5

Alignment Scores:  
Pred. No.: 2.94e-56 Length: 1351  
Score: 606.00 Matches: 120  
Percent Similarity: 62.71% Conservative: 70

Best Local Similarity: 39.60% Mismatches: 111  
 Query Match: 38.43% Indels: 2  
 DB: 4 Gaps: 2

US-10-023-597-24 (1-311) x US-09-524-730-5 (1-1351)

```

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20
DB 149 ATGAAATAGCCATGTAGTTCTCCAGAAAGTCTTCTCCGCGCTTCCGACAGA 208
QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheValAlaValVal 40
DB 209 CCTCCTAGAAAGTCTCTCTTCACTAGTCTCTTCACTAGTCTCTTCACTAGTCT 268
QY 41 GlyAsnLeuGlyLeuIleLeuIleGly--LeuAsnSerArgLeuHisIleProMet 59
DB 269 GGCATAT--GGCATCATCATTTGGTCTCCCATAGATGACATCCACATCCATG 325
QY 60 TyrPhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIlePro 79
DB 326 TACTCTTCTCTGCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 385
QY 80 LysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSer 99
DB 386 CAGCTCCTGCTGCTAAGCTCTGCGGACACAGAAACATTAAGTATGAGAGGTGTG 445
QY 100 GlnPhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetVal 119
DB 446 CAGTCTATATCTCCCATGTGCTGCGGACACAGAAAGTGTCTCTCTCTCTCTCT 505
QY 120 GluAspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGln 139
DB 506 TATGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565
QY 140 ValCysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHis 159
DB 566 CTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 625
QY 160 ThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisIleProMet 179
DB 626 TCCACGCTCACCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 685
QY 180 AspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIleValIle 199
DB 686 GAGATGCCCCCATTAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 745
QY 200 PheIleValIleThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGly 219
DB 746 TACCTGGCCAGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 805
QY 220 PheIleLeuSerSerIleLeuArgValSerSerAlaGluIleArgSerIleAlaPheSer 239
DB 806 CACATTCGCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 865
QY 240 SerCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyr 259
DB 866 ACCGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 925
QY 260 LeuIleProPheSerIleLeuProLeuAspGlnGlyValSerSerLeuPheThrTyr 279
DB 926 CTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 985
QY 280 ThrValValIlePheMetPheAsnProLeuIleTyrSerLeuArgAsnIleValValIle 299
DB 986 GTAGACACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1045
QY 300 AlaLeuIleVal 302
DB 1046 GCCCTCCGG 1054

```

RESULT 10  
 US-09-524-730-5  
 / Sequence 5, Application US/09524730

```

/ Patent No. 6638733
/ GENERAL INFORMATION:
/ APPLICANT: Powers, Scott
/ APPLICANT: Yang, Jianxin
/ APPLICANT: Culler, Gene
/ APPLICANT: Tolarik, Inc.
/ TITLE OF INVENTION: No. 6638733el G-Protein Coupled Receptors
/ FILE REFERENCE: 018781-004710US
/ CURRENT APPLICATION NUMBER: US/09/524,730
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: Patent Ver. 2.1
/ SEQ ID NO 5
/ LENGTH: 1351
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (86) ..(1108)
/ OTHER INFORMATION: human breast cancer amplified G-protein coupled
/ OTHER INFORMATION: receptor 3 (BCA-GPCR-3)
/ US-09-524-730-5

```

#### Alignment Scores:

Score:	2,94e-56	Length:	1351
Percent Similarity:	606.00	Matches:	120
Best Local Similarity:	62.71%	Conservative:	70
Query Match:	38.43%	Mismatches:	111
DB:	4	Indels:	2
		Gaps:	2

US-10-023-597-24 (1-311) x US-09-524-730-5 (1-1351)

```

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20
DB 149 ATGAAATAGCCATGTAGTTCTCCAGAAAGTCTTCTCCGCGCTTCCGACAGA 208
QY 21 ProGlyLeuGlnValProValPhePheLeuGlyPheValAlaValValVal 40
DB 209 CCTCCTAGAAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 268
QY 41 GlyAsnLeuGlyLeuIleIleuIleGly--LeuAsnSerArgLeuHisIleProMet 59
DB 269 GGCATAT--GGCATCATCATTTGGTCTCCCATAGATGACATCCACATCCATG 325
QY 60 TyrPhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIlePro 79
DB 326 TACTCTTCTCTGCGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 385
QY 80 LysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSer 99
DB 386 CAGCTCCTGCTGCTAAGCTCTGCGGACACAGAAACATTAAGTATGAGAGGTGTG 445
QY 100 GlnPhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetVal 119
DB 446 CAGTCTATATCTCCCATGTGCTGCGGACACAGAAAGTGTCTCTCTCTCTCTCT 505
QY 120 GluAspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGln 139
DB 506 TATGACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565
QY 140 ValCysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHis 159
DB 566 CTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 625
QY 160 ThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisIleProMet 179
DB 626 TCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 685
QY 180 AspIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIleValIle 199
DB 686 GAGATGCCCCCATTAATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 745

```



APPLICANT: Jeffrey J. Seilhamer  
 TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
 TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
 NUMBER OF SEQUENCES: 1490  
 CORRESPONDENCE ADDRESS: 1490  
 ADDRESS: INCYTE PHARMACEUTICALS, INC.  
 STREET: 3174 PORTER DRIVE  
 CITY: PALO ALTO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94304  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/016,434  
 FILING DATE: HEREMITH  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0002 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1363:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3459 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: G425220  
 US-09-016-434-1363  
 US-10-023-597-24 (1-311) x US-09-016-434-1363 (1-3459)  
 Alignment Scores:  
 Pred. No.: 4,15e-54 Length: 3459  
 Score: 591.50 Matches: 125  
 Percent Similarity: 58.63% Conservative: 55  
 Best Local Similarity: 40.72% Mismatches: 126  
 Query Match: 37.51% Indels: 1  
 DB: 4 Gaps: 1  
 QY 1 MetAlaAlaGluAsnSerSerSerValIthrgIupheIleuAlaGlyLeuIleHlaGln 20  
 DB 883 ATGATGGAGGCAACCCAGAGTGAAGGTTCCAGAGTTCCTCCGCGGAGTCTCAGAGAGT 942  
 QY 21 ProGlyLeuGlnValIpreValPhePheLeuPheLeuGlyPheYrAlaValIthValVal 40  
 DB 943 CCGTGAAGCAGCAGAGATCTGTTTGGATGTTCTGTCACAGTACCTGTCGACGAGTGTG 1002  
 QY 41 GlyAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArgLeuHlaIleProMetYr 60  
 DB 1003 GGAATGTCCTCATCATCTGCGCATCAGCTCTATTCTCCCGTCGACACCCCGGTATAC 1062  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProGly 80  
 DB 1063 TTCTTCCTGGGCAACCTCTCTTCACAGTACCTCTTCTTGTCAACAACAGATCCCAAG 1122  
 QY 81 MetLeuMetSerPheValSerArgLeuAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 1123 ATGCTGGTGAACCTCCACATCCCATTAACCAAGCCATCTCCATGACAGGAGTCTGACGAG 1182  
 QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120

DB 1183 CTCCTACTCTGCTGCTCTGTCGCGCCCTGGCAACCCATCTGCTGCTGATGCATAT 1242  
 QY 121 AspArgYrValGlyIleCysAsnProLeuLeuYrThrIleThMetSerProGlnVal 140  
 DB 1243 GACCGTATGCGCCATCTGCTGCGCCCTCCCTCAGTACACACAGCCATAGCCCTAAGCTC 1302  
 QY 141 CysLeuLeuLeuLeuGlyValYrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 DB 1303 TGTATCTACTCTTCTTCTGTTGGTGGTCCATCCGCTCTATGAGCTCATACAGACC 1362  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisThrMetCysAsp 180  
 DB 1363 CTCCTCATACCAAGATGACCTTCTGCGGTCACGAAATTCACATCACTTCTGTAG 1422  
 QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerYrIleAsnValIleuValIlePhe 200  
 DB 1423 ATGATATGATTCCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1482  
 QY 201 IleValValIthValIleGlyValIleGlyValIleGlyValIleValIleValIleSerYrGlyPhe 220  
 DB 1483 GCACAGGCTGCTCATCTTCTCTCATCTTCTGATTCGATTCATTCATTCATTCATTCAT 1542  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGlnGlyArgSerYrAlaPheSerSer 240  
 DB 1543 ATTATCAGACCATCTCTCAGATATACCTCAGTCTCTAGAAATACAAAGCCCTTCCACC 1602  
 QY 241 CysSerSerYrIleIleAlaValSerIleuPhePheGlySerGlyAlaPheThrYrIleu 260  
 DB 1603 TGTGCTCCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1662  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheYrThrThr 280  
 DB 1663 AAGCCCTTCATACCTACTCTGTG--AAGACTCAGTACGACACATGATGATGCTGTG 1719  
 QY 281 ValValProMetPheAsnProLeuIleYrSerIleuArgAsnLysAspValIleuAla 300  
 DB 1720 GTGACACCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1779  
 QY 301 LeuLysArgYrThrPheSerArg 307  
 DB 1780 CTGGAGAGACTCTCTAGATPAA 1800  
 RESULT 13  
 US-08-827-291A-1  
 Sequence 1, Application US/08827291A  
 Patent No. 5874243  
 GENERAL INFORMATION:  
 APPLICANT: Macina, Roberto  
 APPLICANT: Sathe, Ganesh  
 TITLE OF INVENTION: NOVEL OLIGOIS RECEPTOR  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: SmithKline Beecham Corporation  
 STREET: 709 Swedeland Road  
 CITY: King of Prussia  
 STATE: PA  
 COUNTRY:  
 ZIP: 19406  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/827,291A  
 FILING DATE: 28-MAR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: King, William T

```

/ REGISTRATION NUMBER: 30,954
/ REFERENCE/DOCKET NUMBER: GP50001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 610-270-5015
/ TELEFAX: 610-270-5090
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1290 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ US-08-827-291A-1

Alignment Scores:
Pred. No.: 1.91e-54 Length: 1290
Score: 589.00 Matches: 118
Percent Similarity: 58.90% Conservative: 64
Best Local Similarity: 38.19% Mismatches: 123
Query Match: 37.35% Indels: 4
DB: Gaps: 1

US-10-023-597-24 (1-311) x US-08-827-291A-1 (1-1290)

QY 1 MetAlaAlaGluAsnSerSerValThrGluPheLeuAlaGlyLeuLeuHisGln 20
Db 296 ATGGCAGGAGAGATTCGACTTCACACTCCGACTTCTTCTGGAATCTTCATCAC 355
QY 21 ProGlyLeuGlnValProValPhePheLeuPheGlyPheValAlaValThrVal 40
Db 356 AGCCCAACCCACACCTTCCTCTTCTGCTGCTCCGACCATCTTCAGTGCCTTCA 415
QY 41 GlyAsnLeuGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 60
Db 416 GGAACCTGTGATGTTCTTCCTCATCTACCTGGACACCCACCTCCACCCCATGAC 475
QY 61 PhePheProPheAsnSerLeuValAspPheSerPheSerThrThrLeuProGly 80
Db 476 CTCCTCCTCAGCAACTGTCCTCCATGACACCTCATGCTATGACACACCGTACCA 535
QY 81 MetLeuMetSerPheValSerValGlyAsnLeuLeuLeuLeuLeuLeuLeuLeu 100
Db 536 ATGGCCTTCAACTGCTGCTGCGACCAAGTCAATTCGCTGCTGCTGCTGCTGCT 595
QY 101 PhePhePhePhePhePhePheValPheSerGluSerPheLeuSerAlaMetValGlu 120
Db 596 ATTTTCTTCTATCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 655
QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
Db 656 GACCGCTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
QY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
Db 716 TGTGACCTATGACTGCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 775
QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
Db 776 GTAGCCACATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 835
QY 181 IleLeuProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 200
Db 836 TTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 895
QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
Db 896 ATTGCTCTATAGTAATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 955
QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGlyGlyArgSerIysAlaPheSerSer 240
Db 956 GTTATCTGCTGCTCATTCATGATGATGATGATGATGATGATGATGATGATGAT 1015

```

```

QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
Db 1016 TGTTCCTCTCAACCCATGAGTGGGATGTTATGTGACAGATTGTTGATGATACATA 1075
QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280
Db 1076 CAGCCACATCTATGCTCTCCCAACGACAGACAGACTGCTGCTGCTGCTGCTGCT 1135
QY 281 ValValProMetPheAsnProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 300
Db 1136 CTCACCTCCATGCTGATCCCTCATCTACAGCTCCGCAACAAGAGAGT----- 1186
QY 301 LeuLysArgThrPheSerArgIleSer 309
Db 1187 ---ACCAAGCATTCATGAAGATCTCA 1210

RESULT 14
US-09-016-434-1313
/ Sequence 1313, Application US/09016434
/ Patent No. 6500938
/ GENERAL INFORMATION:
/ APPLICANT: Janice Au-Young
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
/ NUMBER OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
/ NUMBER OF SEQUENCES: 1490
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/016,434
/ FILING DATE: HEREWITH
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Zeller, Karen J.
/ REGISTRATION NUMBER: 37,071
/ REFERENCE/DOCKET NUMBER: PA-0002 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (650) 855-0555
/ TELEFAX: (650) 845-4166
/ INFORMATION FOR SEQ ID NO: 1313:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1438 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GENBANK
/ CLONE: g32092
/ US-09-016-434-1313

Alignment Scores:
Pred. No.: 6.85e-54 Length: 1438
Score: 584.50 Matches: 115
Percent Similarity: 61.39% Conservative: 71
Best Local Similarity: 37.95% Mismatches: 116
Query Match: 37.06% Indels: 1
DB: Gaps: 1

US-10-023-597-24 (1-311) x US-09-016-434-1313 (1-1438)

```



```

QY 1 MetalalaGluAsnSerSerValThrGlnPheIleuAlaGlyLeuIleHisGln 20
DB ATGAAAGAGAGAACTTACCTTACCTGACTGCTTTGTTCCAGAGTTCTCAGCTC 421
QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrVal 40
DB CATGAGCAGCAGATACCCCTTTGGCGTGTCTCTTGACATACATCTTAACTTGA 481
QY 41 GlyAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
DB GGCAATATCATCATCTGTGATCATCATCGAATATGATCTTCACTTCAACACCACTATC 541
QY 42 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
DB TTCTTCCGAGCAGCTGCTCCACCTTCACTGATGATATACATGATGCTATCTCCACAGA 601
QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
DB ATGCTCTCCAGCCTGTGATGATGAGCCAGCCCATGATGATGAGGAGGTGTCACACAG 661
QY 101 PhePhePhePheCysPhePheValPheSerGlnSerPheIleLeuSerAlaMetValGlu 120
DB ATGTTCTTTTGTGTAACCTTTGGCATCTCACTAAGCTTCTCTGCTCACAGCATGGATAT 721
QY 121 AsparGlyTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
DB GACCCCTATGTCGCCATCTGCAACCCCTGAGATACATGTTATTATGACACAGAGCTG 781
QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisIleThr 160
DB CGATCCCAACTGCTCTGCGGGCCCGCAGCATGGCGATGTAGACATACAGCAAGT 841
QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
DB ACATCTGATTACAGGTACCCCTTCTGTGCT--AGAAAGGTGCCCACTTCTTGTGAC 898
QY 181 IleLeuP-ProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValIleValIlePhe 200
DB ATCCGCCCTGTGATGAAGCTCTCTCGCATGACACACACTGTCATGAATGCTACCTTG 958
QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
DB ATTATCAGTGTGCTGCTGCTGTTGATGCTATGGCTGTGCTTTCATTTCTTATGTTCTC 1018
QY 221 IleLeuSerSerIleLeuValSerValSerAlaGlnGlyArgSerTyrAlaPheSerSer 240
DB ATTATCTTACAACTCTCAAGATGCTTCACTTGAAGGCCGGAAGAGGCTTTTGCCACC 1078
QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
DB TGTGATCCACACTCAGCTGCTGATGTCACATACAGCTGTGCTCATGCTTACCTC 1138
QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerValSerLeuPheTyrThrThr 280
DB AAGCCCAAGTCAGAGACACCAAGACATGATGATCTGGGAGACCTTACCTACCTGTC 1198
QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
DB ATCACTCCCTCCTAGAACCTGTGTGATACACCTGAGAAATTAAGAGGTCAAGATGCT 1258
QY 301 LeuLysArg 303
DB 1259 CTGTGACAG 1267

```

RESULT 15  
 US-09-546-986A-1  
 ; Sequence 1, Application US/09546986A  
 ; Patent No. 6635741  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Powers, Scott  
 ; APPLICANT: Yang, Jianxin  
 ; APPLICANT: Cutler, Gene

```

; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. 6635741el G-Protein Coupled Receptors
; FILE REFERENCE: 018781-004720US
; CURRENT APPLICATION NUMBER: US/09/546,986A
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 09/524,730
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1035
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (42)..(974)
; OTHER INFORMATION: human breast cancer amplified G-protein coupled
; OTHER INFORMATION: receptor 1 (BCA-GPCR-1)
US-09-546-986A-1

Alignment Scores:
Pred. No.: 5,9e-53 Length: 1035
Score: 574.00 Matches: 115
Percent Similarity: 62.33% Conservative: 72
Best Local Similarity: 39.33% Mismatches: 107
Query Match: 36.40% Indels: 6
DB: Gaps: 3

US-10-023-597-24 (1-311) x US-09-546-986A-1 (1-1035)
QY 12 PheIleuAlaGlyLeuIleHisGlnProGlyLeuGlnValProValPhePheLeuPhe 31
DB TTATCTCTTCTGCGGTGTGTGTGACAGCGCGTGTGAGAACCTCTCTTTGTGCTCTC 125
QY 32 LeuGlyPheTyrAlaValThrValAlaGlyAsnLeuGlyLeuIleIleLeuIleGlyLeu 51
DB CTGCTGCTCTATGCTGTGCGCATGTGGGAGACGTCGCATCATCTCCATCCCGGCTG 185
QY 52 AsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSerLeuValAspPhe 71
DB GATCTCAACTCCACAGCCCATGATGATCTTCTCACTGACACTGTCTTCTGACCTC 245
QY 72 SerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLysAsnIle 91
DB TGCTACACACACACAGCATGCTCCAGATGCTGTCACATGAGGAGAGTCCACAGAGCC 305
QY 92 IleSerPheThrGlyCysMetSerGlnPhePhePheCysPhePheValPheSerGln 111
DB ATCAGCTATGAGAGGCTGCACTGTGCAATATGCAATCTTCCACTGCTGGATGCAAGAG 365
QY 112 SerPheIleLeuSerAlaMetValGlnAspArgTyrValGlyIleCysAsnProLeuLeu 131
DB TGCATGCTGCTGCGCCCATGCGCTGAGACCGCTACGTGAGCAGCTGCAACCCCTGAC 425
QY 132 TyrThrIleThrMetSerProGlnValCys-----LeuLeuLeuLeuGlyValTyr 149
DB TATGCGCTTCTCATACACCGCTGCTGTGTAGACAGCTGAGGCTGTGCTGCTGCTGCT 485
QY 150 GlyMetGly---ValPheGlyValAlaAlaHisThrGlyAsnIleValPheLeuThrPhe 168
DB GGCCTTCGGAATCCTTCTGCTGAGAGTGTCTGAGG-----GTCCATTTGCCATTC 536
QY 169 CysAlaAspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGlyLeuSer 188
DB TCGCGCGCGCAGGTGCTGCAACAATTTCTGTGTGAGTGGCGCGCGCTGATCAAGCTGCG 596
QY 189 CysAsnGlySerTyrIleAsnValLeuValIlePheIleValThrValGlyIleGly 208
DB TGTGCTGACACCGCATATGATGACCAATACCTATGCTGTGCTGTGCTTCTTGTGTG 656
QY 209 ValProIleValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgVal 228
DB GTGCCCTGCTCATCTCTCTCTCATGAGCTTATTTGCTCCGCGGAGTGTCTCAGATC 716

```

```

QY      229 SerSerAlaGluGlyArgSerIysAlaPheSerSerCysSerSerTyrIleIleAlaVal 248
Db      717 CAGTCCTCCAAGGAGGACACAAAGGCGCTTTGGAGCGTGTCTCTCCACCTGATGATCGTC 776
QY      249 SerLeuPhePheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeu 268
Db      777 TCCCTCTTCTACCTACCTCGATTACATGTATCTGACGCCCCCTTCCAGCTACTCCCAA 836
QY      269 AspGlnGlyLysValSerSerLeuPheTyrThrThrValValProXerPheAsnProLeu 288
Db      837 GAGCAGGCGCAATTATTCTCTCTCTATTCATATACACCCCACTCTCAATCCCTTC 896
QY      289 IleTyrSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIle 308
Db      897 ACCTACACCCCTGAGAAATAAAGATGAAGGCGGCTGTGAGAGACTTCTGGCCAGGATC 956

```

Search completed: September 30, 2004, 12:13:11  
 Job time : 94 secs

*This Page Blank (uspio)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 09:21:50 ; Search time 2904 Seconds

(without alignments)  
3198.050 Million cell updates/sec

Title: US-10-023-597-24

Perfect score: 1577  
Sequence: 1 MAKENSSSTVEFLAGLHIO.....LRNKVKALKRTFSRISFS 311

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-NODEB=frame+.p2n.model -DEV=xlh  
-Q/cg12\_1/USPTO.spool/US10023597/runac\_29092004.163410.14892/app\_query.fasta\_1.455  
-DB=BSR -CPMT=fastap -SUFPIX=st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -NODE=LOCAL  
-OUTFMT=pro -NORM=ext -HEADSIZE=500 -MITLEN=0 -MAYLEN=2000000000  
-USER=US10023597@CGEN\_1.1.3437@runac\_29092004.163410.14892 -NCFU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.\*  
1: em\_estba.\*  
2: em\_esthm.\*  
3: em\_estln.\*  
4: em\_estlu.\*  
5: em\_estlov.\*  
6: em\_estlp.\*  
7: em\_estro.\*  
8: em\_hic.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hic.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estcom.\*  
17: em\_esthum.\*  
18: em\_estinv.\*  
19: em\_estpin.\*  
20: em\_estvrt.\*  
21: em\_estfun.\*  
22: em\_estmam.\*  
23: em\_estmus.\*  
24: em\_estpro.\*  
25: em\_estrod.\*  
26: em\_estphg.\*  
27: em\_estvrt.\*  
28: gb\_est1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1273.5	80.8	933	29	AY407109 Homo sapi
2	1223.5	77.6	933	29	AY407111 Mus muscu
3	1158	73.4	903	29	AY405611 Mus muscu
4	1139	72.2	903	29	AY405609 Homo sapi
5	1086.5	68.9	799	29	AY407110 Pan trogl
6	1024.5	65.0	930	29	AY405659 Mus muscu
7	958.5	60.8	930	29	AY405657 Homo sapi
8	935	59.3	899	29	AY405610 Pan trogl
9	923.5	58.6	930	29	AY419057 Homo sapi
10	888.5	56.3	930	29	AY419055 Homo sapi
11	880.5	55.8	930	29	AY419056 Pan trogl
12	812	51.5	726	28	AZ255734 RPECI-23-1
13	791	50.2	748	14	CB174479 OR_20E1H1
14	786	49.8	933	29	AY405614 Mus muscu
15	775	49.1	933	29	AY405612 Homo sapi
16	757.5	48.0	848	28	BZ281255 CH230-344
17	724	45.9	945	29	AY402815 Mus muscu
18	710.5	45.1	625	29	CE158635 tigr-gss-
19	707	44.8	733	28	BZ201201 CB174480 OR_20E1H1
20	706	44.8	636	14	CB174480 OR_20E1H1
21	701.5	44.5	942	29	AY402814 Pan trogl
22	700	44.4	637	14	CB174340 OR_2042A0
23	699.5	44.4	945	29	AY402813 Homo sapi
24	686.5	43.5	664	14	CB173618 OR_2021B1
25	684	43.4	924	29	CE184619 tigr-gss-
26	679	43.1	924	29	AY421166 Mus muscu
27	675.5	42.8	752	28	BZ102526 CH230-238
28	669	42.4	897	29	AY405605 Mus muscu
29	664.5	41.7	630	28	AQ0503914 RPECI-11-2
30	657	41.7	936	29	AY405000 Homo sapi
31	655	41.5	936	29	AY405001 Pan trogl
32	654	41.5	939	29	AY405002 Mus muscu
33	651	41.3	897	29	AY405603 Homo sapi
34	650	41.3	939	29	AY405723 Homo sapi
35	650	41.2	945	29	AY404165 Mus muscu
36	649	41.2	907	28	BZ243737 CH230-469
37	648.5	41.1	921	29	AY421164 Homo sapi
38	648	41.1	616	28	BH342053 CH230-51L
39	642	40.7	942	29	AY415251 Homo sapi
40	641	40.6	731	28	BZ207385 CH230-487
41	641	40.6	945	29	AY415715 Mus muscu
42	641	40.6	2021	11	BC016940 Homo sapi
43	640	40.3	1680	11	BC051250 Mus muscu
44	635	40.3	933	29	AY402266 Mus muscu
45	630	39.9	948	29	AY401403 Homo sapi

#### ALIGNMENTS

RESULT 1  
AY407109  
LOCUS AY407109 933 bp DNA linear GSS 15-DEC-2003  
DEFINITION Homo sapiens OREB12 gene, VIRTUAL TRANSCRIPT, partial sequence,  
Genomic survey sequence.  
VERSION AY407109 GI:39763080  
ACCESSION AY407109.1 GI:39763080  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 933)

## AUTHORS

Clark, A.G., Gianowski, S., Nielson, R., Thomas, P., Kejarival, A.,  
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.

## TITLE

Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios

## JOURNAL

Science 302 (5652), 1960-1963 (2003)

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

Clark, A.G., Gianowski, S., Nielson, R., Thomas, P., Kejarival, A.,  
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.

## JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA

## COMMENT

This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.

## FEATURES

Location/Qualifiers

1..933

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

<1..>933

/gene="OR8B12"

/locus\_tag="HCM2781"

## ORIGIN

## Alignment Scores:

Pred. No.: 1,276-98 Length: 933  
 Score: 1273.50 Matches: 250  
 Percent Similarity: 87.46% Conservative: 22  
 Best Local Similarity: 80.39% Mismatches: 38  
 Query Match: 80.75% Indels: 1  
 DB: 29 Gaps: 1

US-10-023-597-24 (1-311) x AY407109 (1-933)

QY 1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20  
 Db 1 ATGGGAGCGCAAAAC--TCTTCTGTGCACAGATTATCTCGAAGGCTTAACCCACAG 57  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheValAlaValThrValAl 40  
 Db 58 CCGGAGCTGGCGATGCCCT 117  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerValGluHisIleProMetTyr 60  
 Db 118 GGGAACTGGGCTTATATACCTGATTGGCTGAACCTCTACCTGCACACTCCCATGTAC 177  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrTrpIleIleProGly 80  
 Db 178 TTCTTCCTTTTAACTCTCTCTTTATATATATTTCTGTTTCTCCATACATCCCAAA 237  
 QY 81 MetLeuMetSerPheValSerValGlyAsnIleIleSerPheThrGlyCysMetSerGln 100  
 Db 238 ATGCTGATGAGATTGTCTCAAGGAACAATCATTTCTTCCACAGGAGTATATGACAG 297  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120  
 Db 298 CTCTTCT 357  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleIleMetSerProGlnVal 140  
 Db 358 GACCGCTACGTGGCCATCTGTACCACTGTTGTACACAGACACCATGCTTGGCAAGTG 417  
 QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 Db 418 TGTTCCT 477  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 Db 478 GGAACATATATGAACTGACCTTCTGTCTGACCAACTTGTCAATCATTCATGTGTGAC 537

QY 181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValIlePhe 200  
 Db 538 ATCT 597  
 QY 201 IleValIleThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 Db 598 ATTGTGTGTCTGTGATGAGTAAAGCCCATGTCACGTCTTATTTATTTATCCCTC 657  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerValAlaPheSerSer 240  
 Db 658 ATCTCTCAGCATTTCTACACAGATTTCTACAGAGGAGGATCCAAAGCCTTAGTACT 717  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyValAlaPheThrTyrIleu 260  
 Db 718 TCCAGTCCACATATATGATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 777  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerValSerSerLeuPheThrThr 280  
 Db 778 AAACCCCTTCCATCTCTGCTCCCTCAGCAGAGGAAGATGCTCTCTCTCTCTATACATA 837  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValIleLeuAla 300  
 Db 838 ATAGTCCCGGTGTAAACCATTAATCTATAGCTTGAGGAACAAGATGCAAGATGCC 897  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 Db 898 CTGAGGAGAACTTGGCGCAAAAAATCTTTTCT 930

RESULT 2  
 AY407111 933 bp DNA linear GSS 15-DEC-2003  
 LOCUS Mus musculus OR8B12 gene, VIRUTAL TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 ACCESSION AY407111.1 GI:39763082  
 VERSION AY407111.1  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 933)  
 Clark, A.G., Gianowski, S., Nielson, R., Thomas, P., Kejarival, A.,  
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE 2 (bases 1 to 933)  
 Clark, A.G., Gianowski, S., Nielson, R., Thomas, P., Kejarival, A.,  
 Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES Location/Qualifiers  
 source 1..933  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>933  
 /gene="OR8B12"  
 /locus\_tag="HCM2781"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,36-94 Length: 933  
 Score: 1223.50 Matches: 237

Percent Similarity: 86.97%  
 Best Local Similarity: 77.20%  
 Query Match: 77.58%  
 DB: 29  
 Gaps: 1

US-10-023-597-24 (1-311) x AY405611 (1-903)

```

QY 1 MetAlaAlaGluAsnSerSerSerValThrGluPheLeuAlaGlyLeuIleHisGln 20
DB 1 ATGACTGCCAGAAAT---TCCTCTGTGATAGAGTTATCTCTGCAGGCTGACAGCCAA 57
QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40
DB 58 CCAGGACTCCGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 117
QY 41 GlyAsnLeuGlyLeuIleLeuIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
DB 118 GGGAACTGGGTTTGATCTCCCTGATGGGCTGAACTTCACCTTCATACCCCTATGTC 177
QY 61 PhePheProPheAsnLeuSerLeuValAlaPhePheSerPheSerThrIleIleProIys 80
DB 178 TCTTTCTCTCTCAATCTTCTCTGTAATAGATTCTGTTATCTCTCACTATCTATCCCAAA 237
QY 81 MetLeuMetSerPheValSerArgLeuAsnIleIleSerPheThrGlyCysMetSerGln 100
DB 238 ATGCTTATGATTTTATCTCAAGAGAACATCATCTCACTCAGGCTGATGACACG 297
QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
DB 298 TGTGTTCTCTCTGTTCTTGTGTTGTTCTGAGTCTCTCATCTGTCAGCAGCATGCAAT 357
QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
DB 358 GACCGTTAGTGGCCATCTGTAACCCCGATGATACAGTACCATGCTCTCAGGTA 417
QY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyValAlaValAlaHisThr 160
DB 418 TATTTACTCTTTTGTCTGGAGCTATGATGGGCTTCTCTGGGCGCATGGCCCATACA 477
QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
DB 478 GGGAACTCATGATGATCTGCTGCTGCTGCAACCTCATCAACATCACTCATCTGAT 537
QY 181 IleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
DB 538 ATCTTCCCTCTCTTACGCTCTCCCTGATAGACCTCATTAAGAGCTGTAATCTTT 597
QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
DB 598 ATGTTGTGGCCCTTGTATGTTGGTGGCCATTTGCACATCTTCTATTCATGCCCTC 657
QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerTyrAlaPheSerSer 240
DB 658 ATCTCTGACGATCTTTCGATGTCATTCACAGAGGAGGTCCAGAGCTTCAGCAC 717
QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
DB 718 TGGAGCTCCACATGATGTTGGTGTGTTCTTTCTTTGTTGGGCTTCATGACTC 777
QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280
DB 778 AAACCTCTTCATTTGCTCTGACCAAGAAAGTCTCTCTGTTCTCATATGAT 837
QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnValAspValIleLeuAla 300
DB 838 GTGCTGCCATGTTGAACCTCTGATCTATGCTTGAAGAAATAGAGATCTCAAGTTCT 897
QY 301 LeuLysArgThrPheSerArg 307
DB 898 CTGAGGAAACCTTGGGTAG 918

```

RESULT 3  
 AY405611

```

LOCUS AY405611 903 bp DNA linear GSS 16-DEC-2003
DEFINITION Mus musculus OREB8 gene, VIRUAL TRANSCRIPT, partial sequence,
ACCESSION AY405611
VERSION AY405611.1 GI:39761585
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 903)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,D.J.,
Adams,M.D. and Cargill,M.
TITLE Interfering nonneutral evolution from human-chimp-mouse orthologous
gene tritos
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 903)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sinsky,D.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
this based on alignment.
FEATURES
location/qualifiers
source
1..903
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>903
/gene="OR8B8"
/locus_tag="HGM2281"
ORIGIN
Alignment Scores:
Pred. No.: 8,45e-89 Length: 903
Score: 1158.00 Matches: 222
Percent Similarity: 84.00% Conservative: 30
Best Local Similarity: 74.00% Mismatches: 48
Query Match: 73.43% Indels: 0
Gaps: 0
DB: 29
US-10-023-597-24 (1-311) x AY405611 (1-903)
QY 12 PheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGlnValProValPhePheLeuPhe 31
DB 1 TTATCTCTTGCAGGCTGACAGACAGCCAGGACTCCGATGCCCTCTCTCTCTCTCT 60
QY 32 LeuGlyPheTyrAlaValThrValGlyAsnLeuGlyLeuIleIleLeuIleGlyLeu 51
DB 61 CTAGCTTTCATACAGTGGAGCTGTGTGGGGAACCTGGGTTGATCTCCCTGATGGGCTG 120
QY 52 AsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSerLeuValAspPhe 71
DB 121 AACTTCACCTCTACACCCCTATGATGATCTTTCTCTTCAATCTTTCTTCAATGATTT 180
QY 72 SerPheSerThrThrIleIleProIysMetLeuMetSerPheValSerArgLysAsnIle 91
DB 181 TGTATCTCTCTACTATCTCCCAAAATGCTGATGATTTTATCTCAAGAAAGACATC 240
QY 92 IleSerPheThrGlyCysMetSerGlnPhePhePheCysPhePheValPheSerGln 111
DB 241 ATCTCACATCCAGGCTGACAGCACATGTTTCTCTCTGTTTCTTGTATCTCTGAG 300
QY 112 SerPheIleLeuSerAlaMetValGluAspArgTyrValGlyIleCysAsnProLeuLeu 131
DB 301 TCTTTATCTGTGAGCAGCATGAGCATATGACCGTTATGTCATCTGTAACCCCGATG 360

```

QY 132 TyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuGlyValTyrGlyMet 151  
 Db 361 TACATGGTCCACCATGCTCCGCCAGTTGGTTGTTACTCTTCTTGGTTGGTATATTATG 420  
 QY 152 GlyValPheGlyValAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAsp 171  
 Db 421 GGTTCGTTGGAGCCATGATGCCACACCATTCATCCATGCGAAGCTGACCTTCTGTGCTGAC 480  
 QY 172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGlnLeuSerCysAsnGly 191  
 Db 481 AACATGTCACACCATATATGATGATGATCTCTCCCTTCTGAGACACCTCTGATCAGC 540  
 QY 192 SerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyValProIle 211  
 Db 541 ACCATGTAAGAACTGAGTCTTATTTTGGATTTGATTTGATTTGGTGTCCCAT 600  
 QY 212 ValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSerSerAla 231  
 Db 601 GTCACCATCTTCATTTCTTATGGCCCTCATCTCTCTACATTTCTCATCATTCACCA 660  
 QY 232 GlnGlyArgSerIleValAlaPheSerSerCysSerSerTyrIleIleAlaValSerLeuPhe 251  
 Db 661 GAGGCGAGCTCCAGCCCTTCAGCACTCCGACCTCCACATGATGTGGTTGTCTTTC 720  
 QY 252 PheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAspGlnGly 271  
 Db 721 TTGCTTCTGGGGCTTTCATGATATCTCAGCCACTTCTGTTCCTTCCCTTGACCAAGA 780  
 QY 272 LysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIleTyrSer 291  
 Db 781 AAGGTGCTCTCTGTTCTATACGATGTGTACTGATGTAATCTCTGATCTATATAGC 840  
 QY 292 LeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSerPheSer 311  
 Db 841 TTGAGAAATTAAGATGTCAAGTACTGTAAGAAACCTTGACAGAGAAATTTTCT 900  
 RESULT 4  
 AY405609 903 bp DNA linear GSS 16-DEC-2003  
 LOCUS Homo sapiens OR8B8 gene, VIRUAL, TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 ACCESSION AY405609.1 GI:39761583  
 VERSION AY405609.1 GI:39761583  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 903)  
 AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Ciavello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trices  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE 2 (bases 1 to 903)  
 AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Ciavello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source location/Qualifiers  
 1..903  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>903  
 gene

ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,516-87  
 Score: 1139.00  
 Percent Similarity: 83.67%  
 Best Local Similarity: 72.33%  
 Query Match: 72.23%  
 DB: 29  
 Gaps: 0  
 US-10-023-597-24 (1-311) x AY405609 (1-903)  
 QY 12 PheIleLeuAlaGlyLeuIleHisGlnProGlyLeuGlnValProValPhePheLeuPhe 31  
 Db 1 TTTATCTCTCGAGGCTTAACTGACCAACCGGAGTCCAGATCCCTCTCTCTCTGTT 60  
 QY 32 LeuGlyPheTyrAlaValThrValValGlyAsnLeuGlyLeuIleIleLeuIleGlyLeu 51  
 Db 61 CTAGGCTTCTAGCTGTGCTACCTGTGTGGGGAACCTGGGCTTGATTAACCTGATAGGCTC 120  
 QY 52 AsnSerArgLeuHisIleProMetTyrPhePheProPheAsnLeuSerLeuValAspPhe 71  
 Db 121 AACTCTCACTTCACACCCCTATGTAATCTTCTCTATTAATCTTGTCTCATATGATTC 180  
 QY 72 SerPheSerThrThrIleIleProLysMetLeuMetSerPheValSerArgLysAsnIle 91  
 Db 181 TGTATATTCAGTGTATATCTCCCAAAATGCTGATGACCTTGTCTTAAAGAAAGAACGC 240  
 QY 92 IleSerPheThrGlyCysMetSerGlnPhePhePheCysPhePheValPheSerGlu 111  
 Db 241 ATCTCTACGACGAGGTGATGACTCAGCTCTTCTTCTTCTTCTTCTTCTTCTTCTGAG 300  
 QY 112 SerPheIleLeuSerAlaMetValGlnAspArgTyrValGlyIleCysAsnProLeuLeu 131  
 Db 301 TCTTCATCTCTCAGCAATGAGCGTACGACCGCTATGTGGCCATTTGTAACCCACTGTG 360  
 QY 132 TyrThrIleThrMetSerProGlnValCysLeuLeuLeuLeuGlyValTyrGlyMet 151  
 Db 361 TACATGGTCCACCATGCTCCGCCAGTTGGTTGTTACTCTTCTTGGTTGGTATATTATG 420  
 QY 152 GlyValPheGlyValAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAsp 171  
 Db 421 GGTTCGTTGGAGCCATGATGCCACACCATTCATCCATGCGAAGCTGACCTTCTGTGCTGAC 480  
 QY 172 AsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGlnLeuSerCysAsnGly 191  
 Db 481 AACATGTCACACCATATATGATGATGATCTCTCCCTTCTGAGACACCTCTGATCAGC 600  
 QY 192 SerTyrIleAsnValLeuValIlePheIleValValThrValGlyIleGlyValProIle 211  
 Db 541 ACCATGTAAGAACTGAGTCTTATTTTGGATTTGATTTGATTTGGTGTCCCAT 660  
 QY 212 ValAlaValPheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSerSerAla 231  
 Db 601 GTCACCATCTTCATTTCTTATGGCCCTCATCTCTCTACATTTCTCATCATTCACCA 660  
 QY 232 GlnGlyArgSerIleValAlaPheSerSerCysSerSerTyrIleIleAlaValSerLeuPhe 251  
 Db 661 GAGGCGAGCTCCAGCCCTTCAGCACTCCGACCTCCACATGATGTGGTTGTCTTTC 720  
 QY 252 PheGlySerGlyAlaPheThrTyrLeuLysProProSerIleLeuProLeuAspGlnGly 271  
 Db 721 TTGCTTCTGGGGCTTTCATGATATCTCAGCCACTTCTGTTCCTTCCCTTGACCAAGA 780  
 QY 272 LysValSerSerLeuPheTyrThrThrValValProMetPheAsnProLeuIleTyrSer 291  
 Db 781 AAGGTGCTCTCTGTTCTATACGATGTGTACTGATGTAATCTCTGATCTATATAGC 840  
 QY 292 LeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSerPheSer 311  
 Db 841 CTGAGAAATTAAGATGTCAAGTACTGTAAGAAACCTTGACAGAGAAATTTTCT 900



[illegible]

ORIGIN /locus\_tag="HCM2297"

Alignment Scores:  
 Pred. No.: 2,02e-77 Length: 930  
 Score: 1024.50 Matches: 193  
 Percent Similarity: 77.74% Conservative: 48  
 Best Local Similarity: 62.26% Mismatches: 68  
 Query Match: 64.97% Indels: 1  
 Gaps: 1

US-10-023-597-24 (1-311) x AY405657 (1-930)

```

OY 1 Metalaalagluasenserserservalthrgupheileuialeuilehisln 20
Db 1 ATGGCTCCAGAAAGAGCTCTCTGTGACTGATTTATCTTGCGGATTTTCAACCA 60
OY 21 ProglyleuiglinalprovalphepheleupheleuglypneThyralavalTh 40
Db 61 CCAGCTCCAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
OY 41 Glyaenleuglyleuileleleuileleuileuansersaragluahisilep 60
Db 121 GGCAATTTGGAGCTGATCACTTAATGGGTTGANTTCACTTCACTTCACTGATAC 180
OY 61 Phepheprophasnleuenserleuvalaspheserserthrtthleleprolys 80
Db 181 TTTTCTCTTCAACTTCTCTCTTAATGATTTCTGTATCTCTGTGTTCCAGCCAAA 240
OY 81 MetleuMetSerPheValSerarglysaenlleleuSerPheThnglyCysMetSer 100
Db 241 ATGTGAGTGAATTTGTCTCA--GAAACATTAATCTCTCAAGGAGGATGACTGAA 297
OY 101 PhephephephecysPhePheValPheSerGlySerleuSerAlaMetValglu 120
Db 298 CTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 357
OY 121 AsparglytyrvalglylleCysasnProleuLeuTyrrthlethMetSerProgl 140
Db 358 GATCCCTATGTGGCCATTGCAATCTCTGTGTAACAAGATACCATGTCCTCCAGGC 417
OY 141 CysleuLeuLeuLeuGlyValTyrglyMetGlyValPheGlyAlaValAlaHis 160
Db 418 TGCATCTGCTGATCTTTTGTTCATATGTGATAGGTTTGTGAGGCAATGCCCACT 477
OY 161 GlyasnleuValPheLeuThrPheCysAlaAspAsnleuValAsnHisTyrrMe 180
Db 478 GGAAACATGCTGACACTTACTTCTGTATTCACATATGATCCACATTAATCTCTG 537
OY 181 IleleuProleuLeuLeuSerCysasnGlySerTyrrleasnValleuValIle 200
Db 538 GTTCTTCCCTTTTACAGCTCTCTCTGCAACAGTACTTATGCCAAGAGCTTTG 597
OY 201 IleValAlaThValGlylleGlyValProleuValAlaPheIleSerTyrgly 220
Db 598 ATTGTTGTTGAGTATGATCATCAGACATCCAGCATCTTATCTTATGACATG 657
OY 221 IleleuSerSerIleleuArgValSerSerIleGlyArgSerIleAlaPheSer 240
Db 658 ATCTCTCCACATTTTAAGATCTCTCTCTCTCTGAGGATAGATCAAAAGCTTT 717
OY 241 CysSerSerTyrrleleuAlaValSerleuPhePheGlySerGlyAlaPheTh 260
Db 718 TGTGTTCCCATGATGATGCTGCTGCTGTTTGGTGAGGAGGATCACTATTTA 777
OY 261 LysPProPheSerIleleuProleuAspGlnGlyValSerSerleuPheThTh 280
Db 778 ACACCTCTTCCCTGCTCAATGGAAGAGGAGGAGGTTGCTCAGATTTTATAC 837
OY 281 ValValPrometPheAsnProleuIleTyrrSerleuArgAsnIleAspVal 300
Db 838 GTGGTTCCCATGCTTAACCACTGATCTATAGTTTGAGATTAAGAGCTTAAC 897

```

OY 301 LeuTyArgThrPheSerArgIleSerPhe 310  
 Db 898 CTTAAATAAACCTTGAAGAGAGTCTCTTC 927

RESULT 7  
 AY405657  
 LOCUS  
 DEFINITION  
 Homo sapiens OR8B4 gene, VIRUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 AY405657  
 ACCESSION  
 AY405657.1 GI:39761631  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 930)  
 Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanendrum,D.M., Civiello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Adams,M.D. and Cargill,M.  
 inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios

TITLE  
 JOURNAL  
 Science 302 (5652), 1960-1963 (2003)  
 PUBMED  
 14671302  
 2 (bases 1 to 930)

REFERENCE  
 AUTHORS  
 Clark,A.G., Gianowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanendrum,D.M., Civiello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA

COMMENT  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.

FEATURES  
 source  
 1..930  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>930  
 /gene="OR8B4"  
 /locus\_tag="HCM2297"

gene

ORIGIN

Alignment Scores:  
 Pred. No.: 8.42e-72 Length: 930  
 Score: 958.50 Matches: 186  
 Percent Similarity: 74.84% Conservative: 46  
 Best Local Similarity: 60.00% Mismatches: 77  
 Query Match: 60.78% Indels: 1  
 Gaps: 1

US-10-023-597-24 (1-311) x AY405657 (1-930)

```

OY 1 Metalaalagluasenserserservalthrgupheileuialeuilehisln 20
Db 1 ATGGCTCCAGAAAGAGCTCTCTGTGACTGATTTATCTTGCGGATTTTCAACCA 60
OY 21 ProglyleuiglinalprovalphepheleupheleuglypneThyralavalTh 40
Db 61 CCAGGCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
OY 41 Glyaenleuglyleuileleleuileleuileuansersaragluahisilep 60
Db 121 GGCAATTTGGAGCTGATCACTTAATGGGTTGANTTCACTTCACTTCACTGATAC 180
OY 61 Phepheprophasnleuenserleuvalaspheserserthrtthleleprolys 80
Db 181 TTTTCTCTTCAACTTCTCTCTTAATGATTTCTGTATCTCTGTGTTTCCAGCCAAA 240
OY 81 MetleuMetSerPheValSerarglysaenlleleuSerPheThnglyCysMetSer 100

```

Db	24	ATGCGAATGACTTGTSTTCA---GAAAGATCATCTCTTACTGCGATGACTCAG	297
Qy	101	PhePhePhePheCysPhePheValPheSerGluSerPheIleuSerAlaMetValGlu	120
Db	298	CTATTTTCTTCTGTTCTTGTTCATTTCTGATGCTATGTTGGTATCAATGAGCCCTAT	357
Qy	121	AspArgIValGlyIleCysAsnProLeuLeuThrIlePheMetSerProGlnVal	140
Db	358	GATCGCTATGTGGCCATCTGCACACCCCTCTCTATATGCTACCATGTCCTCCAGAGTC	417
Qy	141	CysLeuLeuLeuLeuGlyValIYrgIYmetGlyValPheGlyValAlaAlaHisThr	160
Db	418	TGCTTTCGCTGATGTTGGTTCCTATGTGATGAGAGGTTTGCTGGGGCCATGGCCCACT	477
Qy	161	GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisIYrMetCysAsp	180
Db	478	GGAAGCATCTCGATGACCTTCCTGTGATTCACAGTCATGACCATTAATCTGTGTGAC	537
Qy	181	IleLeuProLeuLeuGluLeuSerCysAsnGlySerIYrIleAsnValLeuValIlePhe	200
Db	538	GTTCTCCCTCTGTGAGCTCTCTCTCCACAGACCCATGTCAGTACGCTGATTTTTC	597
Qy	201	IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerIYrGlyPhe	220
Db	558	ATTGTTGTGGAGTATACCATCATGCTATCCAGCATGATGATGTCATCTCTTAAGCTTTG	657
Qy	221	IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerIYrValIlePheSer	240
Db	658	ATACTCTCCACATCTCTCTGATCTCTTCTGACAGAGGACAGATCCAAACCTTTAGCA	717
Qy	241	CysSerSerIYrIleIleAlaValSerLeuPheGlySerGlyAlaIleThrIYrLeu	260
Db	718	TGGGCTCCCACATATTCCTGTTGCTGTTTGGTGTGACGAGCATTCACCTACTTA	777
Qy	261	LysProProSerIleLeuProLeuAspGlnGlyIYsValSerSerLeuPheIYrThr	280
Db	778	ACAAATCTTTTCTCGGCTCTTAAACCATGACAGATTGCTTCAGTCTTTTAAACCAT	837
Qy	281	ValValPrometPheAsnProLeuIleIYrSerLeuArgAsnIYsAspValIYsLeuAla	300
Db	838	GTGGTTCCTCATCTTAACCTTCGATCTACAGTTTGAGGAATTAAGATGATTAACCTGCC	897
Qy	301	LeuIYsArgThrPheSerArgIleSerPhe	310
Db	898	CTGGCGCAAAACCTTAAGAGAGTCTCTTC	927
RESULT 8			
AY405610	899 bp	DNA	linear
LOCUS	Pan troglodytes OR888 gene, VIRTUAL TRANSCRIPT, partial sequence,		GSS 16-DEC-2003
DEFINITION	genomic survey sequence.		
ACCESSION	AY405610		
VERSION	AY405610.1		
KEYWORDS	GSS.		
ORGANISM	Pan troglodytes (chimpanzee)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.		
	1 (bases 1 to 899)		
	Clark,M.A.G., Gianowski,S., Nielson,R., Thomas,P., Kejariwal,A.,		
	Todd,M.A., Tenenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,		
	Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,		
	Adams,M.D. and Cargill,M.		
TITLE	Interfing nonneutral evolution from human-chimp-mouse orthologous		
JOURNAL	gene erios		
PUBMED	Science 302 (5552), 1960-1963 (2003)		
REFERENCES	14671302		
AUTHORS	2 (bases 1 to 899)		
	Clark,M.A.G., Gianowski,S., Nielson,R., Thomas,P., Kejariwal,A.,		
	Todd,M.A., Tenenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,		
	Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,		
	Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		

JOURNAL Submitted (16-Nov-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source 1. .899 /organism="Pan troglodytes" /mol\_type="genomic DNA" /db\_xref="taxon:9598" <1. .2899 /gene="OR8B8" /locus\_tag="HCM2281"

ORIGIN

Alignment Scores:

Pred. No.:	8.19e-70	Length:	899
Score:	935.00	Matches:	184
Percent Similarity:	72.91%	Conservative:	34
Best Local Similarity:	61.54%	Mismatches:	81
Query Match:	59.29%	Indels:	0
DB:	29	Gaps:	0

US-10-023-597-24 (1-311) x AY405610 (1-899)

```

QY      12  Phe11eLeuA1aG1yLeu1eH1sG1nProG1yLeuG1nValProValPhePheLeuPhe 31
Db      1  TTTATCTCTGACAGGCTTAACTGACCAACCGGAGTCCAGATCCCCCTCTTCTCTGTTT 60

QY      32  LeuG1yPhe1yValA1aThrValValG1yAsnLeuG1yLeu1e1Leu1eG1yLeu 51
Db      61  CTAGGCTTCACTGGATCTGACTGTGGGGGAACTGNNNNNNNAACNNATAGTCTTC 120

QY      52  AsnSerArgLeuHis1e1eProMetC1yPhePheProPheAsnLeuSerLeuValAspPhe 71
Db      121 AACNNTCACNNACACACCCCTNNNAACNTCTTCTCTTAACCTTGTCTTCAATGATTTTC 180

QY      72  SerPheSerThrThr1e1e1eProLysMetLeuMetSerPheValSerArgLysAsn1e 91
Db      181 TGCTATATCCANNNTTATCACTCCCAAAATGCTGAGACTTGTCTTANNNNNNACAGC 240

QY      92  11SerPheThrG1yCyMetSerG1nPhePhePhePheCySphPheValPheSerG1u 111
Db      241 ATCTCTACGACAGGNNNTATGACTCANNNTCTTACTCTTCTTCTTGTGTCTGTG 300

QY      112 SerPhe1eLeuSer1aMetValG1uAspArg1yValG1y1e1CyAsnProLeuLeu 131
Db      301 TCTCTCATCTCTGACCAATGCGATATGACANNNTATGTGGCATCTGTAAACCACTGTTC 360

QY      132 TyrThr1e1eH1eMetSerProG1nValCyS1eLeuLeuLeuLeuG1yVal1yG1yMet 151
Db      361 TACATGATGATACATGNNNCCACAGGTGTGTTTCTCCCTTTGTGGGNNNNATGGGATG 420

QY      152 G1yValPheG1yValA1aVal1aH1sThrG1yAsn1eValPheLeuThrPheCyAlaAsp 171
Db      421 GGGTTGGCTGGGGCCATGAGCCCAACACAGCGATGATGAGGTGTGACCTNNNGCCAAAT 480

QY      172 AsnLeuValAsnHis1yTrMetCyAsp1eLeuProLeuLeuG1uLeuSerCySAsnG1y 191
Db      481 AACCTGTCAACACACTACATG1NNNACATCCCTCCCTCANNAGTGGCTTGACACAGC 540

QY      192 SerTyr11eAsnValLeuVal1ePhe1e1eVal1aThrVal1aG1y1eG1yValPro1e 211
Db      541 ACCTAATGTAAAGAGCTTGATGTGTGTGTGTGTGTGGCATTTGATGTATGTGTGGTCCACA 600

QY      212 ValA1aValPhe1eSerTyrG1yPhe1eLeuSerSer1eLeuArgValSerSerAla 231
Db      601 GTCAACATCTTCA1TCTCATATCTCTCAATCTCANNACATCTTCCACATGATTCACAG 660

QY      232 GluG1yArgSerLysValAlaPheSerSerCySerSerTyr11e1e1aValSerLeuPhe 251
Db      661 GAGGGGAGNTCCAAAGCTTCAAGCATCGACGCTCCCAATATGTGAGTTTCTNNNC 720

QY      252 PheG1ySerG1yAlaPheThrTyrLeuSerProSerSer1eLeuProLeuAspG1nG1y 271
  
```



PUBMED 14671302  
 REFERENCE 2 (bases 1 to 930)  
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snijsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT These sequences were made by sequencing genomic exons and ordering  
 them based on alignment  
 FEATURES  
 source  
 1..930  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>930  
 /gene="OR8A1"  
 /locus\_tag="HCM6763"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7,698-66 Length: 930  
 Score: 888.50 Matches: 172  
 Percent Similarity: 74.678 Conservative: 55  
 Best Local Similarity: 56.588 Mismatches: 76  
 Query Match: 56.344 Indels: 1  
 DB: 29 Gaps: 1  
 US-10-023-597-24 (1-311) x AY419055 (1-930)  
 QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleLeuAlaGlyLeuIleIleGln 20  
 DB 1 ATGGCTGCGAGAAATACCTCTACAGAGACAGAGGTCATCTCAAGGTTTACGAGAGA 60  
 QY 21 ProGlyLeuGlyValProValPhePheLeuGlyLeuGlyPheTyrAlaValThrValAl 40  
 DB 61 GCAGACCTCCAGCTCCCT 120  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuIleIleProMetTyr 60  
 DB 121 GGGAACCTGGGACATGACCTAAATTGTGTGAAGCTCAGCTGACACCCCATGTAC 180  
 QY 61 PhePheProPheAsnLeuSerLeuValAlaPheSerPheSerThrIleIleProIys 80  
 DB 181 TACTTCTCAGACATGTGACATGATCTCTGCTACTCTCTGCTCATTAACCCCTAAG 240  
 QY 81 MetLeuMetSerPheValSerArgLeuAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 241 ATGCTGTGACATCTGTCTGAGAAACATCATCTCTCTACGAGGGGTGATGTACAG 300  
 QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120  
 DB 301 CTCTACT 360  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140  
 DB 361 GACCGTATGTTGCCATCTGCCACCTTTGCTTTACACATCATTAATGTCTATCAGACC 420  
 QY 141 CysLeuLeuLeuLeuGlyValIleTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 DB 421 TGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAlaHisThrMetCysAsp 180  
 DB 481 GGGCTCATGTAAATGCTCTATTTGTT--GAGACCTCATATGATCAGCTCTGTGAC 537  
 QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAspValLeuValIlePhe 200  
 DB 538 ATCTCTCCCTCATGAAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 597  
 QY 201 IleValValThrValGlyIleGlyValIleProIleValAlaValPheIleSerTyrGlyPhe 220

DB 598 TTTTGGCTGGATTCAACATCATAGTCACGAGCTTAACAGTCTCTGTTCTTACACCTTC 657  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGlyIleArgSerIleValAlaPheSerSer 240  
 DB 658 ATTCTCTCCAGCATCTCGGACATGACACACAGAGGGAGATCCAAAGCCTTCAGCACC 717  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPheGlySerGlyAlaPheThrTyrLeu 260  
 DB 718 TGCAGCTCCACACTTGACCGGTGGAGATTTCTATGATCATCACTGATCATGTACTTA 777  
 QY 261 LysProProSerIleLeuProLeuAspGlyIleValSerSerLeuPheTyrThrThr 280  
 DB 778 AAACCTCCACAAATCACTTCTTACCCACAGAGATGTGCTGTGTCTACACCACG 837  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnIleValIleValAla 300  
 DB 838 GTATATCCCATGTTGAATCCCTTAATCTACAGCTGAGAACAGAAAGAGCTGCTGCC 897  
 QY 301 LeuLysArgThr 304  
 DB 898 GTCCAGAAAACG 909  
 RESULT 11  
 AY419056 930 bp DNA linear GSS 17-DEC-2003  
 LOCUS  
 DEFINITION Pan troglodytes OR8A1 gene, VIRtual TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 ACCESSION AY419056  
 VERSION AY419056.1 GI:3975016  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.  
 REFERENCE  
 1 (bases 1 to 930)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Snijsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT These sequences were made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source  
 1..930  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 <1..>930  
 /gene="OR8A1"  
 /locus\_tag="HCM6763"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,698-65 Length: 930  
 Score: 880.50 Matches: 171  
 Percent Similarity: 74.34% Conservative: 55  
 Best Local Similarity: 56.25% Mismatches: 77  
 Query Match: 55.83% Indels: 1  
 DB: 29 Gaps: 1  
 US-10-023-597-24 (1-311) x AY419056 (1-930)

```

OY 1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20
DB 1 ATGGCTGAGAGAAATCACTTACAGTAGCAGAAATTCATCTCAAGGTTAANNAGAGA 60
OY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPhePheValAlaValThrValVal 40
DB 61 GCAAGCCTCCAGCCTCCCTCTTCTCTCTCTCCGGAGTCACTAGTGGTCAACATCATG 120
OY 41 GlyAlaLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
DB 121 GGGAACTCGGGCATGATCATCTTAATTTGTGTGAACCTCGACGACACCCCATGTAC 180
OY PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLys 80
DB 181 TACTTCCCTCAGCAATCTCTCACTGCTGATCTCTGCTACTCTCTCCGCTTATACCCCAAG 240
OY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
DB 241 ATGCTGCTGAACTTGTGTGAGAGAAAACATCTCTCTCCGACGAGGTGATGTCACAG 300
OY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120
DB 301 CTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
OY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
DB 361 GACGCTATGTTGCTGATGCTGACCTTGTCTTACAACATCATTAATGCTCATACACC 420
OY 141 CysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyValAlaValAlaHisThr 160
DB 421 TGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
OY 161 GlyAlaIleValPheLeuThrPheCysAlaAspAsnLeuValAsnIleTyrMetCysAsp 180
DB 481 GGCCTCATGTTAAACCTCCCTATTGT--GACGACCTTATCACTCACTCTCTCTGAC 537
OY 181 IleLeuProLeuLeuLeuLeuSerCysAsnGlySerTyrIleAsnValIleValIlePhe 200
DB 538 ATCTCTCTCTCATGAAAGCTGTCTCTGCTGACCTCATGATGTTGAGAGACAGCTTCC 597
OY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
DB 598 TTTTGGCTGATTCGACATCATAGTCACGAGCTTAACAGTCTTCTTCTTACACCTTC 657
OY 221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerLysAlaPheSerSer 240
DB 658 ATTCTCTCCAGCATCTCGGCATCAGACACCAAGAGGGAGATCCAAAGCCTTCAGACCC 717
OY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
DB 718 TGCACTTCCACCTTCGACCGCTGGAGATGTTCTATGATCACTGATTCATGACTTCA 777
OY 261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThrThr 280
DB 778 AAACCTTCACCAATCATGTTCTTGAACCCAGAGAAATGTGCTGTGTTCTACACACG 837
OY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300
DB 838 GTATATCCCATGTTGAATCCCTTATATCAAGCTGAGAACAAAGAAAGAGGCTGCC 897
OY 301 LeuLysArgThr 304
DB 898 GTGCAAGAAAACG 909

```

```

RESULT 12
A2255734 726 bp DNA linear GSS 26-JUL-2000
LOCUS RPECI-23-16619.TU RPECI-23 Mus musculus genomic clone RPECI-23-16619,
DEFINITION genomic survey sequence.
ACCESSION A2255734
VERSION A2255734.1 GI:9458784
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)

```

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 726)
REFERENCE
Zhao,S., Nierman,W., Feldblum,T., Malek,J., Shatman,S.,
Akinret,B., Levins,M., McGann,S., Tesgaye,G., Geer,K., Krol,M., de
Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPECI-23
JOURNAL
Unpublished (1995)
Other_GSS: RPECI-23-16619.TV
COMMENT
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@igr.org
Clones are derived from the mouse BAC library RPECI-23. For BAC
library availability, please contact Pieter de Jong
(pjeter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Reseach Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 166 Row: I Column: 9
Seq primer: SP6
Classes: BAC ends.
FEATURES
source
1..726
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPECI-23-16619"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPECI-23"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1:
EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBAC3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
ORIGIN
Alignment Scores:
Pred. No.: 2,026-59 Length: 726
Score: 812.00 Matches: 149
Percent Similarity: 78.01% Conservative: 39
Best Local Similarity: 61.83% Mismatches: 53
Query Match: 51.49% Indels: 0
DB: 28 Gaps: 0
US-10-023-597-24 (1-311) x A2255734 (1-726)
OY 25 ValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValAlaGlyAsnLeuGly 44
DB 2 CTCCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 61
OY 45 LeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyrPhePheProPhe 64
DB 62 TTGGTAATTTCTAATTTGAGACTGAATTCACCTTCACACCCCGAGTACTTTTCTCTTT 121
OY 65 AsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProLysMetLeuMetSer 84
DB 122 AACTGTGTTGATGACCTCTGTACTCTTCACTGTTTACACCCAAAGCTGTGAAAC 181
OY 85 PheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGlnPhePhePhe 104
DB 182 TTCATATTAATAATTAATATATCTTATATGCGGGGTATAGACCACTATTTTAT 241
OY 105 CysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlyAspArgTyrVal 124

```





Genomic survey sequence.  
 AY405614  
 VERSION  
 AY405614.1 GI:39761588  
 KEYWORDS  
 GSS.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 1 (bases 1 to 933)  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,  
 Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Interfering nonneutral evolution from human-chimp mouse orthologous  
 gene tris  
 TITLE  
 Science 302 (5652), 1960-1963 (2003)  
 JOURNAL  
 PUBMED  
 14671302  
 REFERENCE  
 2 (bases 1 to 933)  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,  
 Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source  
 1..933  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>933  
 /gene="ORS5A1"  
 /locus\_tag="HGM2282"  
 ORIGIN  
 Alignment Scores: 4.12e-57 Length: 933  
 Pred. No.: 786.00 Matches: 151  
 Score: 68.40% Conservative: 59  
 Percent Similarity: 49.19% Mismatches: 97  
 Best Local Similarity: 49.84% Indels: 0  
 Query Match: 29 Gaps: 0  
 US-10-023-597-24 (1-311) x AY405614 (1-933)  
 QY 1 MetAlaAlaGluAsnSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20  
 DB 1 ATGATTAAGAAATCACTCAGTGTGATGAATTTGCTTTATGGGCATCACTCAAGAC 60  
 QY 21 ProGluGluGlnValProValPhePheLeuGlyPheTyraIValThrValVal 40  
 DB 61 CCTCAGCTGCAAGATCATTTTGTGTGCTCTCTCTGTCATCCTGCTCATGTGATA 120  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 DB 121 GCGAATGTTGATGATCATCTCATATACAGACAGCTTCAGCCCATGTAC 180  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrTrilleProLys 80  
 DB 181 TTTTCTCTGCAACCTCTCTCTGTTGACCTGGGCTACTCATAGCTATGCTCCAGG 240  
 QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 241 ATGCGGCTGATTTTCTTAACAAGACAAAGTATATCTTTTCCACGTGGCCACCCAG 300  
 QY 101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120  
 DB 301 TTGCTTCTTGTAGAGCTTTGTGATGCTAGGCTAGGCTGCGGCTGAATGGCCTAT 360  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140

DB 361 GACCGCTTGGGCCATTCGACACCTCTTCACTATAGCACTTCATGTCACAAAGATC 420  
 QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 DB 421 TGCTTGGCTTAAGCTGGCTGCTTACTCTGCTGCTAGGATGATTAGGCGCATACT 480  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspLeuValAsnHisTyrMetCysAsp 180  
 DB 481 TCTCTCACTTCAGCTCAGTACTGTGCTTCCATATCAACCAACCATTTCTTGTGGA 540  
 QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValIleuValIlePhe 200  
 DB 541 ATCCACACACTTATAGCTCTCTCTGTCAGATACCTATAGATGAGACTCTGCTTT 600  
 QY 201 IleValIleThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 DB 601 AGTTGTGTGCTTCATGATGATTCAGACACCTCTTATCATCTTCATCTCTATGCCCTTC 660  
 QY 221 IleLeuSerSerIleLeuArgValSerSerArgIleGlyArgSerIleAlaPheSerSer 240  
 DB 661 ATTTCTATTCATCATCATCATGATTCATGATTCATGATTCATGATTCATGATTCATGATTC 720  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260  
 DB 721 TGCGGCTCTCATCTTACTGCTGTATACCTTTTATGTTATGATGATGATGATGATGATGAT 780  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheTyrThr 280  
 DB 781 AGGCCAATCATCATCTCTCTCTGACCAACAGAGTGGCTCTGCTTTCATCAAT 840  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300  
 DB 841 ATCATCTCTATGCTCAACCCCTTGATTCACAGTTTGGAGAACAGATGAAGCCGCT 900  
 QY 301 LeuLysArgThrPheSerArg 307  
 DB 901 TTCAAGAAACTGATTGGAAG 921  
 RESULT 15  
 AY405612 933 bp DNA linear GSS 16-DEC-2003  
 LOCUS  
 DEFINITION  
 Homo sapiens ORS5A1 gene, VIRUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 AY405612  
 ACCESSION  
 AY405612.1 GI:39761586  
 VERSION  
 KEYWORDS  
 GSS.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 933)  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,  
 Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Interfering nonneutral evolution from human-chimp mouse orthologous  
 gene tris  
 TITLE  
 Science 302 (5652), 1960-1963 (2003)  
 JOURNAL  
 PUBMED  
 14671302  
 REFERENCE  
 2 (bases 1 to 933)  
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civeille, D.R., Lu, F., Murphy, B.,  
 Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT  
 This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 source  
 1..933  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"

gene  
/db\_xref="taxon:9606"  
<1..>933  
/gene="OR5A1"  
/locus\_tag="HCM2282"

## ORIGIN

## Alignment Scores:

Pred. No.:	3,56e-56	Length:	933
Score:	775.00	Matches:	151
Percent Similarity:	67.64%	Conservative:	58
Best Local Similarity:	48.87%	Mismatches:	100
Query Match:	49.14%	Indels:	0
DB:	29	Gaps:	0

US-10-023-597-24 (1-311) x AY405612 (1-933)

```

QY      1 MetAlaAlaGluAsnSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20
        |||
        1 ATGGATAAAGAAAGACGCTCATGTGACTGAGTTATCTTCATGGGCATCACCGAGAC 60
QY      21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyAlaValThrVal 40
        |||
        61 CCTCAGATGAGATCATCTTCGTCGCTTCCTCATAGTTACTGCTGTTATGATGAGT 120
QY      41 GlyAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60
        |||
        121 GGGAAATATGTGATGATGATTCCTGATTACACAGACACTCAGCTTCACACACCATGAT 180
QY      61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleProLys 80
        |||
        181 TTTTCTCTGCACCTCTCTTTGTTGACCTGGCTACTCCTCAGCCATGCCCCCAG 240
QY      81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
        |||
        241 ATGCTGCTGACTTCTTAACAAATCAAAAGTTATCTCTTCACAGCTGTGCCACCG 300
QY      101 PhePhePhePheCysPhePheValPheSerGluSerPheIleLeuSerAlaMetValGlu 120
        |||
        301 TTGCTTTTGTGAGTTTGTGAGTGTGAGTGTGATGCTGCTGCGACGATGGCTAT 360
QY      121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140
        |||
        361 GGTGTTTGTGGCCTTTGTGACCCCTCCACTATAGCACCCTTCATGCCAAGCAGTTC 420
QY      141 CysLeuLeuLeuLeuGlnValTyrGlyMetGlyValPheGlyValAlaValAlaHisThr 160
        |||
        421 TGCTTGCTCTCATGCTGGCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY      161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180
        |||
        481 ACCCTCACCCTGAGCTGAGTTACTGTGCTTCCAAATATCATCATCATTTCTTCGCAA 540
QY      181 IleLeuProLeuGlnLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200
        |||
        541 ATCCACCACTCTTGCCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY      201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220
        |||
        601 AGTCTGTGGCTTATGAATTCAGCACCATCTCATCTTCATCTTCATCTTCATCTTC 660
QY      221 IleLeuSerSerIleLeuArgValSerSerAlaGlnGlyArgSerIleValAlaPheSerSer 240
        |||
        661 ATCTTGTGCAATCATCAGAAATGCGTTAGCTGAAGCGCGCTTAAGGCTTCTCCACC 720
QY      241 CysSerSerTyrIleIleValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260
        |||
        721 TGGGGTCTCACCTTACCTGACATCACCTCTTCATGACACATCATGTTATGACTCTG 780
QY      261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280
        |||
        781 AGGCCAACAATCAGCTACTCCCTGGACCAAGTGGCGCTCTGTGTCTACACAGGTT 840
QY      281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300

```

```

DB      841 ATCATCCCATGTAAATCCCTTGATCTACAGTTTGGGAACAAGATGGAAGCTGCT 900
QY      301 LeuLysArgThrPheSerArgIleSer 309
        |||
DB      901 TTCAAAAGCTAATTGAAAAAATCT 927

```

Search completed: September 30, 2004, 12:11:38  
Job time : 2912 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 12:13:16 ; Search time 4423 Seconds

(without alignments)  
3047.634 Million cell updates/sec

Title: US-10-023-597-24

Perfect score: 311

Sequence: 1 MAENSSVTEFLAGLIHQ.....LKNKVKLAKRTFSRISFS 311

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516935 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934735

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-O/can2.1/USPTO.spool/US10023597/runat.29092004.163430.15410/app\_query.fasta.1.455  
-DB=GenEmbl -QFMT=fastap -SUFFIX=olig.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZE=500 -MINLEN=2000000000  
-USER=US10023597 @CCN 1.1 3731 @runat.29092004.163430.15410 -NCPU=6 -ICPU=3  
-NO MMAP -LARGECUTERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.\*  
1: gb\_ba:.\*  
2: gb\_hlg:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_sts:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*  
15: em\_ba:.\*  
16: em\_fun:.\*  
17: em\_hum:.\*  
18: em\_in:.\*  
19: em\_mu:.\*  
20: em\_om:.\*  
21: em\_or:.\*  
22: em\_ov:.\*  
23: em\_pat:.\*  
24: em\_ph:.\*  
25: em\_pl:.\*  
26: em\_ro:.\*  
27: em\_sts:.\*  
28: em\_un:.\*

29: em\_vi:.\*  
30: em\_htg\_hum:.\*  
31: em\_htg\_inv:.\*  
32: em\_htg\_other:.\*  
33: em\_htg\_mus:.\*  
34: em\_htg\_pln:.\*  
35: em\_htg\_rtd:.\*  
36: em\_htg\_mam:.\*  
37: em\_htg\_vrt:.\*  
38: em\_sy:.\*  
39: em\_htgo\_hum:.\*  
40: em\_htgo\_mus:.\*  
41: em\_htgo\_other:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	* Match Length	DB ID	Description
1	311	100.0	6 AX555964	AX555964 Sequence
2	210	67.5	6 AX241500	AX241500 Sequence
3	210	67.5	2 AC083958	AC083958 Homo sapi
4	210	67.5	9 AP000916	AP000916 Homo sapi
5	191	51.4	6 AX646877	AX646877 Sequence
6	105	33.8	6 AY448475	AY448475 Gorilla g
7	74	23.8	6 AR427386	AR427386 Sequence
8	74	23.8	6 BD122939	BD122939 EST and e
9	70	22.5	6 AY448924	AY448924 Cercopit
10	50	16.1	6 AY448357	AY448357 Hylobates
11	41	13.2	6 AY448759	AY448759 Cercopit
12	40	12.9	6 AX646875	AX646875 Sequence
13	40	12.9	6 AX241493	AX241493 Sequence
14	34	10.9	6 AY449293	AY449293 Saimiri s
15	33	10.6	6 AX181578	AX181578 Sequence
16	33	10.6	6 AF179842	AF179842 Saimiri b
17	28	9.0	6 AY448836	AY448836 Callithrix
18	23	7.4	6 AY448503	AY448503 Alouatta
19	21	6.8	6 AY448669	AY448669 Eulemur m
20	21	6.8	6 AY449228	AY449228 Ateles fu
21	21	6.8	9 AY073346	AY073346 Mus muscu
22	21	6.8	10 AY318057	AY318057 Mus muscu
23	21	6.8	2 AC131161	AC131161 Rattus no
24	21	6.8	2 AC097099	AC097099 Rattus no
25	21	6.8	2 AC096935	AC096935 Rattus no
26	21	6.8	2 AX181444	AX181444 Sequence
27	20	6.4	6 AF179764	AF179764 Homo sapi
28	20	6.4	6 AF399511	AF399511 Homo sapi
29	20	6.4	6 AX242246	AX242246 Sequence
30	20	6.4	10 AF282281	AF282281 Mus muscu
31	20	6.4	6 AX244609	AX244609 Sequence
32	20	6.4	6 AX448443	AX448443 Sequence
33	20	6.4	6 BD144473	BD144473 Novel G-P
34	20	6.4	10 AY073208	AY073208 Mus muscu
35	20	6.4	6 AX554459	AX554459 Sequence
36	20	6.4	6 AX556206	AX556206 Sequence
37	20	6.4	6 AX702842	AX702842 Sequence
38	20	6.4	6 AX646099	AX646099 Sequence
39	20	6.4	6 AB065834	AB065834 Homo sapi
40	20	6.4	6 AX357037	AX357037 Sequence
41	20	6.4	6 AP001524	AP001524 Homo sapi
42	20	6.4	2 AC083958	AC083958 Homo sapi
43	20	6.4	2 AC109178	AC109178 Mus muscu
44	20	6.4	2 AC131161	AC131161 Rattus no
45	20	6.4	2 AC131161	AC131161 Rattus no

RESULT 1

#### ALIGNMENTS

AX555964  
 LOCUS AX555964 963 bp DNA linear PAT 27-NOV-2002  
 DEFINITION Sequence 23 from Patent WO0250275.  
 ACCESSION AX555964  
 VERSION AX555964.1 GI:25899399  
 KEYWORDS

SOURCE  
 ORGANISM Homo sapiens (human)

Homosapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS Padigaru, M., Kekuda, R., Li, L., Ballinger, R.A., Casman, S.J.,  
 Spytek, R.A., Baumgartner, J.C. and Burgess, C.E.

TITLE Novel proteins and nucleic acids encoding same  
 JOURNAL Patent: WO 0250275-A 23 27-JUN-2002;  
 Curagen Corporation (US)

FEATURES  
 source Location/Qualifiers  
 1..963  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

## Alignment Scores:

Pred. No.:	1..376-316	Length:	963
Score:	311.00	Matches:	311
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
Ds:	6	Gaps:	0

US-10-023-597-24 (1-311) x AX555964 (1-963)

QY 1 MetAlaAlaGluAsnSerSerSerValThrgluPheIleuAlaGlyLeuIleHisGln 20  
 Db 2 ATGGCTGCGAGAACTCTCTCCGTGACAGATTATCCCGAGGTTATATCCACAG 61  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThValVal 40  
 Db 62 CCGGAGCTCCAGGCTCCGCTCTTCTCTTCTAGGTTTCTACGGCGTCACGCTG 121  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 Db 122 GGGAACTCGGCTTATATCTCTGATAGGCTCACTCCGCTGCATATCCCATATAC 181  
 QY 61 PhePheProPheAsnLeuSerIleuValAspPheSerPheSerThrThrIleIleProLys 80  
 Db 182 TTTTCCCTTCACTGCTCCCTCGTAGATTAGTTCTCTACGACCATCATCCCAA 241  
 QY 81 MetLeuMetSerPheValSerArgIleAsnIleIleSerPheThrGlyCysMetSerGln 100  
 Db 242 ATGCTGATGAGTTTGTCTCAAGAGAACTATTTCTTCTCAAGGCTGATAGTCAG 301  
 QY 101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120  
 Db 302 TTTCTTCTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTGTTCTG 361  
 QY 121 AspArgIleValGlyIleCysAsnProLeuLeuTyrThrIleIleMetSerProGlnVal 140  
 Db 362 GACCCCTACGTGGGATCTGTAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 421  
 QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 Db 422 TGTTCCTCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnIleuValAsnHisTyrMetCysAsp 180  
 Db 482 GGAATATATGTTCTTCTCACTTTTGTGACAGCAACTTGTCAATCACTACATGTGAC 541  
 QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValIleValIlePhe 200  
 Db 542 ATCTTCCCTTCTTCTGAGCTCTCTGCAAGGCTCTTACATTAATGCTCTGCTATCTT 601

QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 Db 602 ATTGTTGTACCGCTTGGGATGGGCTGCCATTTGTTGCTTTTATATCTTATGCTTTT 661  
 QY 221 IleLeuSerSerIleLeuArgValSerSerAlaGluGlyArgSerValAlaPheSerSer 240  
 Db 662 ATCTTCTTCAGCATCTCCGGTTAGTTCTCTGAGGCGCAGCTCAAGCCCTCAGTAC 721  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrIleu 260  
 Db 722 TCGAGCTCTCATATATATGCACTTTCTTTTCTTTGCTGAGGAGCTTTTACGACTTC 781  
 QY 261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280  
 Db 782 AAACCCCTTCCATTTTACCCCTGACCAAGGAGAAAGTCTCTCTGTTCTATACCACT 841  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerIleuArgAsnLysAspValIleLeuAla 300  
 Db 842 GTGGTGCCATTTTATACCATTTATCTACAGCCCTGAGAGATTAAGATGTCAACTTCC 901  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 Db 902 CTGAAGAGAACCTTTTCCAGATTAAGCTTTTCT 934  
 RESULT 2  
 AX241500  
 LOCUS AX241500 994 bp DNA linear PAT 26-SEP-2001  
 DEFINITION Sequence 248 from Patent WO0127158.  
 ACCESSION AX241500  
 VERSION AX241500.1 GI:15798375  
 KEYWORDS  
 SOURCE Synthetic construct  
 ORGANISM synthetic construct  
 ORGANISM artificial sequences.  
 REFERENCE  
 AUTHORS 1 Bellenson, J., Smith, D., Lancel, D., Gusman, G., Fuchs, T. and  
 Yanai, I.  
 TITLE Olfactory receptor sequences  
 JOURNAL Patent: WO 0127158-A 248 19-APR-2001;  
 DISCLOSURE (US): YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)  
 Location/Qualifiers  
 source 1..994  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="(H3897 nucleotide)"  
 ORIGIN  
 Alignment Scores:  

Pred. No.:	1..72e-210	Length:	994
Score:	210.00	Matches:	310
Percent Similarity:	99.36%	Conservative:	0
Best Local Similarity:	99.36%	Mismatches:	1
Query Match:	67.52%	Indels:	2
Ds:	6	Gaps:	0

US-10-023-597-24 (1-311) x AX241500 (1-994)

QY 1 MetAlaAlaGluAsnSerSerSerValThrgluPheIleuAlaGlyLeuIleHisGln 20  
 Db 1 ATGGCTGCGAGAACTCTCTCCGTGACAGATTATCCCGAGGTTATATCCACAG 60  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThValVal 40  
 Db 61 CCGGAGCTCCAGGCTCCGCTCTTCTCTGTTCTAGGTTTCTACGGCTCACGGTGTG 120  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 Db 121 GGGAACTCGGCTTATATCTGATAGGCTCACTCCGCTGCATATCCCATATAC 180  
 QY 61 PhePheProPheAsnLeuSerIleuValAspPheSerPheSerThrThrIleIleProLys 80  
 Db 181 TTTTCCCTTCACTGCTCCCTCGTAGATTAGTTTCTCTACGACCATCATTCCTCAA 240

```

QY      81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100
Db      241 ATCTGATGATGTTTGTCTCAAGGAAAGACATATTTCCTTCAAGGCTGTATGACTCG 300
QY      101 PhePhePhePheCysPhePheValPheSerGlySerPheIleLeuSerAlaMetValGlu 120
Db      301 TTTCTTCTTCTTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 359
QY      121 -AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThMetSerProGlnTyr 140
Db      360 GGACCCCTACGTCGGGATCTGTATACCCACTGTTGTACAGATCACATGTCTCCCAAGT 419
QY      140 LysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160
Db      420 GTGTTTGTCTTCTTCTTCTGAGTCTCTACGGGATGCGGGGTTTGGGGCTGTGCTCATAC 479
QY      160 rgiYAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAs 180
Db      480 AGGAAATATAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 539
QY      180 PileuLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePh 200
Db      540 CATCTTCCCTCTTCTTCTGAGTCTCTGCAACGGCTCTTACATTAATGCTCTGCTCATCT 599
QY      200 eileValIleValIleGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPh 220
Db      600 TATGTTGTGACCGGTTGGGATGGGGTGGCCATGTGTGCGGTTTATCTTATAGTGT 659
QY      220 eileLeuSerSerIleLeuArgValSerSerAlaGlyIleArgSerTyrIleAlaPheSerSe 240
Db      660 TATCTTCCAGCATCTCTCCGCTTGTCTGTGAGGCGAGCTCTTAAGCTTCACTG 719
QY      240 rCysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyValAlaPheThrTyrLe 260
Db      720 CTGCACTCTCTCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 779
QY      260 uLysPProPProSerIleLeuProLeuAspGlnGlyLysValSerSerLeuPheThrTh 280
Db      780 CAACCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 839
QY      280 rValValPProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValIleLeuAl 300
Db      840 TGTGATGCCCATGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 899
QY      300 AleuLysArgThrPheSerArgIleSerPheSer 311
Db      900 CCTGAAGAGAACCTTTTCCAGATAAGCTTTTCT 933

RESULT 3
AC083958
LOCUS   AC083958               172991 bp    DNA          linear   HTG 03-MAR-2001
DEFINITION Homo sapiens clone RP11-18M5, WORKING DRAFT SEQUENCE, 8 unordered
          pieces.
AC083958
VERSION AC083958.2   GI:13.84080
HTG: HTGS PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS  Biren,B., Linton,L., Nusbaum,C. and Lander,E.
          1 (bases 1 to 172991)
          Homo sapiens chromosome, clone RP11-18M5
          Unpublished
          2 (bases 1 to 172991)
          Biren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
          Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
          Chongel,Y., Colangelo,M., Collins,S., Collipalmo,A., Cooke,P.,
          Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
          Fitzhugh,W., Gage,D., Galagan,J., Gardys,S., Glade,S., Goyette,M.,
          Graham,L., Grand-Pierre,N., Hags,B., Hatford,A., Horton,L.,

```

# TITLE JOURNAL COMMENT

Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Labrecque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G.,  
Macdonald,P., Margulis,N., McCarthy,M., McEwan,P., McKernan,K.,  
McNeerers,R., Meldrim,J., Menue,L., Mihova,T., Mlenga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neill,D., Olivar,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sougnuez,C., Spencer,N., Strange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tefaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliou,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.

Direct Submission  
Submitted (08-OCT-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Mar 2, 2001 this sequence version replaced gi:10717225.  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WtBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
Project Information  
Center project name: L11149  
Center clone name: 18.M.5

Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye, 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 169811 bases at least Q40  
Consensus quality: 170912 bases at least Q30  
Consensus quality: 171602 bases at least Q20  
Insert size: 176000; agarose-fp  
Insert size: 172291; sum-of-contigs  
Quality coverage: 7.3 in Q20 bases; agarose-fp  
Quality coverage: 7.4 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently  
consists of 8 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 388: contig of 388 bp in length  
\* 389 488: gap of 100 bp  
\* 489 3383: contig of 2895 bp in length  
\* 3384 3483: gap of 100 bp  
\* 3484 6775: contig of 3292 bp in length  
\* 6776 6875: gap of 100 bp  
\* 6876 69914: contig of 63039 bp in length  
\* 69915 70014: gap of 100 bp  
\* 70015 87665: contig of 17651 bp in length  
\* 87666 87665: gap of 100 bp  
\* 87666 112075: contig of 24310 bp in length  
\* 112076 112175: gap of 100 bp  
\* 112176 166387: contig of 54212 bp in length  
\* 166388 166487: gap of 100 bp  
\* 166488 172991: contig of 6504 bp in length.

## FEATURES SOURCE

1..172991  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-18M5"  
/clone\_1fb="RP11-18M5"  
/clone\_1fb="RP11-18M5"  
/note="assembly\_firmament  
1..388  
clone\_end:SP6

```
misc_feature      489.3333333333333 vector_side:left"
misc_feature      3484.6775 /note="assembly_fragment"
misc_feature      6876.65914 /note="assembly_fragment"
misc_feature      70015.87655 /note="assembly_fragment"
misc_feature      87766.112075 /note="assembly_fragment"
misc_feature      112176.165387 /note="assembly_fragment"
misc_feature      166488.172991 /note="assembly_fragment"
misc_feature      clone_end:17 vector_side:right"
```

Alignment Scores:	
Pred. No.:	1,56e-308
Score:	210.00
Percent Similarity:	99.36%
Best Local Similarity:	99.36%
Query Match:	67.52%
DB:	2
	Gaps: 0
	Length: 17299
	Matches: 30
	Conservative: 0
	Mismatches: 1
	Indels: 0
	Gaps: 0

US-10-023-597-24 (1-311) X AC083958 (1-172991)

Oy	1	MethylalaguaneserSerValThGlPheIleuValIleuValIleuHisGln	20
Dd	94318	ATGGCTCCGAAGAACTCCTCCTCGTAGCAAGAATTATTCCTGCAGGCTTATTCACAG	94377
Oy	21	ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheThrAlaValThrValVal	40
Dd	94378	CCGGGACTCCAGGCTCCCGCTTCTTCCTGTTCGTAGGTTTCTAACCGGCTACCGGGTG	94437
Oy	41	GlyAsnuGlyLeuIleIleIleuIleGlyLeuAsnSerArgLeuHisIleProMetTyr	60
Dd	94438	GGGAACCTGGGCTTGATTAATCTGATAGGAGCTCACTCTGGCTCGATATCCCAGTAC	94497
Oy	61	PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleLeuPolys	80
Dd	94498	TTTTTCCCCCTCAACTGTCCCTCGTAGATTTAAGTTTCTTAGGCCATCATTTCCCAA	94557
Oy	81	MetIeuMetSerPheValIserArgLysAsnIleIleSerPheThrIlyCyMetSerGln	100
Dd	94558	ATGCTGATGAATTTGTCCTACAAGGAACAATTAATTCCTTCACAGSGTGATGAGTCAG	94617
Oy	101	PhePhePhePheCySPhePheValPheSerGluSerPheIleuSerXlaMetValGlu	120
Dd	94618	TTCCT	94676
Oy	121	-AspArgTyValGlyIleCySaenProLeuLeuTyrThrIleTh-MetSerProGlnVa	140
Dd	94677	GGACCGCTACCTGGGCACTGTATAACCACACTGTTGTACACATCACCATGTCTCCAGGT	94736
Oy	140	LysLeuLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisIsm	160
Dd	94737	GTTGTTTGCTCCTTTTACTGGGTGTCTMGCGGATGGGGGTTTTGGGGCTGTGGCTATAC	94796
Oy	160	rgLYAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAs	180
Dd	94797	AGGAAAATAATGTTTCTCACTTTTGTGAGCAACAATTGTAAATCACACTACATGTGTGA	94856
Oy	180	pIleLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePh	200
Dd	94857	CATCCTTCCTCTTTAGAGCTCTCTGCAAGGCTCTTAATAAAGTCTCGTGCATCTT	94916
Oy	200	eIleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPh	220
Dd	94917	TATGTGTGTGACCGTGGCATTTGGAGGCCAATGTGTGCGTTTATCTTATATGTTT	94976

Oy		220	eileusSerSeriIeLeuArgValSerSerIaGIlyARSerYsaIabSeSe	240
Dd		94977	TATCTTTCACGACTTCGCCGGTAGTTCTGTGCAGCGCAGGCTCAAGCCITCAGTAG	95036
Oy		240	rCySeSerSYrYIlelleAvalSerIeuPhePhgIySerIylabheThryrIte	260
Dd		95037	CCTGACCCTCTCAATAATTGACGTTTCTCTTTCTTGCGCACAGACCTTTACGAACCT	95096
Oy		260	uLySPcProSoSerIleLeuProLeuAspGLngIyLysValSerSerDeuPheYrThrh	280
Dd		95097	CAAAACCCCTTCATTTAACCCCTGAGCACAGGGAAAGTGCCTCCCTGTTCTATACAC	95156
Oy		280	rValVLPrometPheashProLeuIIeTYrSerIeuHArGsmIysaspIallYlsIeuAl	300
Dd		95157	TGTGTGCCCAAGTTAACCCATTATATCTTAACGCTGACGAATAAGATGTCAAACCTGC	95216
Oy		300	aLeuLYArGrTHrPheSerArGlSeSePheSer	311
Dd		95217	CCTGAAGAAGACCTTTTCCAGATAAGCTTTTCT	95250
RESULT 4				
AP000916/c		AP000916	192926 bp	DNA linear PRI 01-FEB-2003
LOCUS		Homo sapiens genomic DNA, chromosome 11 clone:RP11-728D14, complete sequence.		
DEFINITION		AP000916		
ACCESSION		AF000916.6	GI:28189528	
VERSION		AF000916		
KEYWORDS		HTG.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS		1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki.Y.		
TITLE		2 Published Only in Databse (1999)		
JOURNAL		2 (bases 1 to 192926)		
REFERENCE		Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki.Y.		
AUTHORS		Direct Submission		
TITLE		Submitted (17-DEC-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suenho-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170]		
JOURNAL		On Jan 31, 2003 this version replaced gi:15637083.		
COMMENT		Location/Qualifiers		
FEATURES		source		
		1..192926		
		/organism="Homo sapiens"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:9606"		
		/chromosome="11"		
		/map="11q"		
		/clone="RP11-728D14"		
ORIGIN				
Alignment Scores:				
Fred. No.:		1..72e-208	Length:	192926
Score:		210.00	Matches:	310
Percent Similarity:		99.36%	Mismatch:	0
Best Local Similarity:		99.36%	Mismatches:	1
Query Match:		67.52%	Indels:	2
DB:		9	Gaps:	0
US-10-023-597-24 (1-311) x AP000916 (1-192926)				
Oy		1	MetAlaAlaGIuAnSerSeSerValThrgIupHeilleuAlaGIyLeuIIehISglN	20
Dd		102700	ATGGCTGGCGAGAACTCTCTCCGAGACAGATTATCTCCGACGCTTATCATCACAG	102641
Oy		21	PrrGlyIeuGlInValProValPhePheIeuPheIeuGlyPheTyfAlaValThValVal	40



Db	102644	CCGGGACCTCCAGGATCCCGCTCTCTTCCTTCGTTTCTTAGTTCTTAGCGGTTTACCGCGGTCACGGTGGTG	102581
QY	41	GIYAsnLeuGIYLeuIleIleLeuIleGIYLeuAsnSerArgLeuHISLeuProMetTYR	60
Db	102580	GGAAACCTGGGGCTGTATAATCCGATAGAGGAGCTCAACTTCGCGCTGGATATACCCCAATGTAC	102521
QY	61	PhePheProPheAsnLeuSerLeuValIAspPheSerPheSerThrThrIleLeuProlys	80
Db	102520	TTTTTCCCTTCACTGTGTCCTCGTAAAGTTTAGTTTCTCTACGACATCATCTCCAAA	102461
QY	81	MetLeuMetSerPheValSerArgIYAsnIleIleSerPheThrGIYMetSerGln	100
Db	102460	AGCTGATGAGTTTGTCTCCAGGAAGAACATTAATTCCTTCACAGGAGTATAGACTCAG	102401
QY	101	PhePhePhePheCysPhePheValPheSerGlnSerPheIleLeuSerIAspMetValGlu	120
Db	102400	TTCTTCTTCTTCTGTGTCTTCTTGTCTTCTTCTGACTCTTCATCTCTGCGGAGTGGTGA	102342
QY	121	AspArgTYRValGIYIleCysAsnProLeuLeuTYRThrIleThrMetSerProGlnVa	140
Db	102341	GGACCGCTACGTGGGCATCTGTAAACCACTGTGTGTCAAGATCACCATGCTCTCCAGG	102282
QY	140	LysLeuLeuLeuLeuLeuGIYValTYRGIYMetGIYValPheGIYAlaValAlaHisThr	160
Db	102281	GGTTTGTCTCTTTTACTGGGTGTCAAGGGAATGGGAGTTTGGGAGTGGCTCATATC	102222
QY	160	IGIYAsnIleValPheLeuThrPheCysIAspAsnLeuValAsnHISTYRMetCysAs	180
Db	102221	AGGAATATATAGTGTCTCTACCTTTGTGTGAGACAACTTGTCAATCACTACATGTGTGA	102162
QY	180	PileuProLeuLeuGIYLeuSerCysAsnGIYSerTYRLeuAsnValLeuValIlePh	200
Db	102161	CATCTTCCCTCTTGTAGACTCTCCGCAACGCGCTTTCATATAAAGTCTGTCTCATCTT	102102
QY	200	eIleValValThrValGIYIleGIYValProIleValAlaValPheIleSerTYRGIYPh	220
Db	102101	TATTTGTGTGACCGTGGCATTTGGGATGGCCCATGTGTGCCGTTTATCTCTTAGGGTTT	102042
QY	220	eIleLeuSerSerIleLeuArgValSerSerAlaGlnGIYArgSerLYsAlaPheSerSe	240
Db	102041	TATTTCTTTCACACATCTCCGCGTGTGTCTGTGAGGACAGGCTTAAGCCTTCAGTAC	101982
QY	240	TYSerSerSerTYRleIleAlaValSerLeuPhePheGIYSerGIYAlaPheThrTYRle	260
Db	101981	CTGCAGCTCTTACATATATGCAAGTTCCTCTTCTTGAGTCAGAGCTTTTACGAACT	101922
QY	260	ULYsProProSerIleLeuProLeuAspGlnGIYLYsValIleSerSerLeuPheTYRThr	280
Db	101921	CAAAACCCCTTCATTTTACCCCTGTGACCAAGGAGAAAGTCTCTCTCTGTATATCCAC	101862
QY	280	TYAlaValProMetPheAsnProLeuIleTYRSerLeuAlaGlnLYsAspValLYsLeuAl	300
Db	101861	TGTGTGCCCAATGTTTAAACCACTTATCTACAGCCGAGAGAAATTAAGATGTCAAACTTGC	101802
QY	300	aLeuLYsArgThrPheSerArgIleSerPheSer	311
Db	101801	CCTGAAGAGAACCTTTTCCAGATTAAGCTTTTCT	101768
RESULT 5			
LOCUS	AX646877	931 bp	DNA linear PAT 04-MAR-2003
DEFINITION	Sequence 1069 from Patent EP1270724.		
ACCESSION	AX646877		
VERSION	AX646877.1	GI:28799404	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	Suwa,M., Asai,K., Akiyama,Y and Aburatani,H.		
JOURNAL	Guanosine triphosphate-binding protein coupled receptors		
	Patent: EP 1270724-A 1069 02-JAN-2003;		

FEATURES		Location/Qualifiers	
source		1..931	
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
CDS		201..731	
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAD69565.1"		
	/db_xref="GI:28799405"		
	/db_xref="REMBL:CAD69565"		
	/translation="MSPOVCLILLGVYGVGVGAVAHNTNIVFLTFCADNLVNHVM DLPRLSLSCNSGYINIVVIRIVTVVGIQPIYAVPISYFILLSLIRPSAGREK FSCSSYIIAIVSLFFPGSAFTYLNKPSILPLDQKVSIFPTTVVEFNPILYSNN DVKLAKRFRSIRISFS"		
ORIGIN			
Alignment Scores:			
Pred. No.:	1,51e-190	Length:	931
Score:	19.00	Matches:	241
Percent Similarity:	99.18	Conservative:	0
Best Local Similarity:	99.18	Mismatches:	1
Query Match:	61.41%	Indels:	2
DB:	6	Gaps:	0
US-10-023-597-24 (1-311) x AM646877 (1-931)			
QY	70	AspPheSerPheSerThrThrIleLeProlysMetLeuMetSerPheValSerAllys	89
DB	3	GATTTTAGTTCTCTACGACCATCATCTCCAAATGCGATGATTTTGTCTCAAGAG	62
QY	90	AsnIleIleSerPheTherGlyCysMetSerGlnPhePhePheCysPhePheValPpe	109
DB	63	AACTATTATTCCTTCAAGAGGTATGAGTCAATTCCTTCTTCTTCTTCTTCTTCTT	122
QY	110	SerGlySerPheIleuSerAlaMetValGlu-AspArgIyValGlyIleCysAnPr	129
DB	123	TCTGAGCCTTCATCCTGCGCGAGATGTGA-GAACCGCTAGCGGACATCTGTAACC	181
QY	129	oleuLeuTyThrIleThMetSerProGlnValCysIleuIleuIleuIleuValIy	149
DB	182	ACGTGTGACACATACCATCATCTCCACAGGTGTGTCTCTTCTTCTTCTTCTTCTA	241
QY	149	rgIyMetGlyValPheGlyAlaValAlaIleThGlyAsnIleValPheLeuThrPheCy	169
DB	242	CGGAGTGGGGGTTTTGGGCGGTGGCGTCACACACAGAAATATAGTTTCTCACCTTTG	301
QY	169	SalAspAsnLeuValAsnHisTyrMetCysAspIleuProLeuIleuGlyLeuSerCy	189
DB	302	TGAGAGCAACTGTGCAACACTACATGTGAGATCTCTCCCTTGTAGCTCTCCG	361
QY	189	saenGlySerTyIleAsnValIleValIlePheIleValIleValThrValGlyIleGlyVa	209
DB	362	CAACGGCTTACATTAATATCTGTGTATCTTTATTTGTTGACCGTTGGCATTTGGGCT	421
QY	209	IProIleValAlaValPheIleSerTyGlyPheIleuSerSerIleLeuArgValSe	229
DB	422	GCCCATTTGTGCGGTTTATCTCTTAGGTATTATCTTCCAGCATTTCCCGCTTAG	481
QY	229	rSerAlaGlyArgSerIyAlaPheSerCysSerSerCysSerSerTyIleIleAlaValSe	249
DB	482	TTCGTGAGGGAGGCTTAAGCCTTAGAGCTGACGCTCTACATTAATTCAGATTTC	541
QY	249	rLeuPhePheGlySerGlyAlaPheThrTyIleuIleProProSerIleLeuProLeuS	269
DB	542	TCTTTTCTTTGGGTGAGAGCTTTTACGTACTCAAAACCCCTTCATTTTACCCCTGGA	601
QY	269	pgInGlyIyValSerSerIePheTyThrThrValValProMetPheAsnProIleuI	289
DB	602	CCAGGAGAAAGTCTCTCCCTGTCTTAAACAATGTGAGTCCCAATGTTTAAACCATTAAT	661

QY 289 eTyzSerLeuArgAsnLysAspValLysLeuAlaLeuLysArgThrPheSerArgIleSe 309  
DB 662 CTACAGCCTGAGGAATTAAGATGTCATTAACCTTGACGAGAACCTTTCACAAATATAG 721

QY 309 rPheSer 311  
DB 722 CTTTCT 728

## RESULT 6

AY448475 667 bp DNA linear PRI 07-DEC-2003  
LOCUS AY448475  
DEFINITION Gorilla gorilla clone OLG\_68 olfactory receptor-like protein gene,  
partial sequence.

ACCESSION AY448475  
VERSION AY448475.1 GI:38634503  
KEYWORDS  
SOURCE  
ORGANISM Gorilla gorilla (gorilla)  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Gorilla.

REFERENCE  
AUTHORS Gilad, Y., Wiebe, V., Przeworski, M., Lancel, D. and Paabo, S.  
TITLE Loss of olfactory receptor genes coupled with the acquisition of  
full trichromatic vision in primates

JOURNAL  
REFERENCE  
AUTHORS Gilad, Y., Wiebe, V., Przeworski, M., Lancel, D. and Paabo, S.  
TITLE Direct Submission  
JOURNAL Submitted (20-OCT-2003) Max Plank Institute for Evolutionary  
Anthropology, Deutscher Platz 6, Leipzig 04103, Germany

FEATURES  
Source location/Qualifiers

1..667  
/organism="Gorilla gorilla"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9593"  
/clone="OLG 68"  
/note="sample obtained from Primate Genetics German  
Primate Center in Goettingen, Germany"  
1..>667  
/gene="olfactory receptor-like protein"

## ORIGIN

Alignment Scores:  
Pred. No.: 2.73e-100 Length: 667  
Score: 105.00 Matches: 105  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 33.76% Indels: 0  
DB: 9 Gaps: 0

US-10-023-597-24 (1-311) x AY448475 (1-667)

QY 135 ThrMetSerProGlnValCysLeuLeuLeuLeuGluValTyrGlyMetGlyValPhe 154  
DB 215 ACCATGTCCTCCAGAGTGTTGCTTACTTTTACGGGAGTCACGGGATGGGGTTTTT 274  
QY 155 GtAlaValAlaAlaHisThrGlyAsnLysValPheLeuThrPheCysAlaAspAsnLeuVal 174  
DB 275 GGGGCTGTGGCTCATACAGAAATATAGTTTTCACCTTTTGGGAGACACCTTGTTC 334  
QY 175 AsnHisTyrMetCysAspLysLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIle 194  
DB 335 AATGCTACTAGTGTGACATCTTCCCTTCTTGAGCTCTCTGCAAGGCTCTTACATA 394  
QY 195 AsnValLeuValIlePheIleValValThrValGlyIleGlyValProIleValAlaVal 214  
DB 395 AATGCTACTAGTGTGACATCTTCTTGAGCTCTCTGCAAGGCTCTTACATA 454  
QY 215 PheIleSerTyrGlyPheIleLeuSerSerIleLeuArgValSerSerAlaGlyIleArg 234  
DB 455 TTATCTCTTATGATGTTTATCTTTCACGACATTCCTCCGCTAGTTCTGCTGAGGAGG 514

QY 235 SerLysAlaPheSer 239  
DB 515 TCTAAAGCCTTCAGT 529

## RESULT 7

AR427386/c 485 bp DNA linear PAT 18-DEC-2003  
LOCUS AR427386  
DEFINITION Sequence 18883 from patent US 6639063.  
ACCESSION AR427386  
VERSION AR427386.1 GI:40182496  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE  
AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.  
TITLE EST's and encoded human proteins  
JOURNAL Patent: US 6639063-A 18883 28-OCT-2003;  
Location/Qualifiers

1..485  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

Alignment Scores:  
Pred. No.: 7.85e-68 Length: 485  
Score: 74.00 Matches: 74  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 23.79% Indels: 0  
DB: 6 Gaps: 0

US-10-023-597-24 (1-311) x AR427386 (1-485)

QY 135 ThrMetSerProGlnValCysLeuLeuLeuLeuGluValTyrGlyMetGlyValPhe 154  
DB 263 ACCATGTCCTCCAGAGTGTTGCTTACTTTTACGGGAGTCACGGGATGGGGTTTTT 204

QY 155 GtAlaValAlaAlaHisThrGlyAsnLysValPheLeuThrPheCysAlaAspAsnLeuVal 174  
DB 203 GGGGCTGTGGCTCATACAGAAATATAGTTTTCACCTTTTGGGAGACACCTTGTTC 144

QY 175 AsnHisTyrMetCysAspLysLeuProLeuLeuGluLeuSerCysAsnGlySerTyrIle 194  
DB 143 AATGCTACTAGTGTGACATCTTCCCTTCTTGAGCTCTCTGCAAGGCTCTTACATA 84

QY 195 AsnValLeuValIlePheIleValValThrValGlyIleGly 208  
DB 83 AATGCTCTGTCATCTTATTTGTTGTGACCGTTGCACTTGGG 42

## RESULT 8

BD122939/c 485 bp DNA linear PAT 18-SEP-2002  
LOCUS BD122939  
DEFINITION EST and encoded human protein.

ACCESSION BD122939  
VERSION BD122939.1 GI:23217864  
KEYWORDS JP 2002010789-A/15016.  
SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.  
TITLE EST and encoded human protein  
JOURNAL Patent: JP 2002010789-A 15016 15-JAN-2002;  
GENEST CORP

## COMMENT

OS Homo sapiens (human)  
PN JP 2002010789-A/15016  
PD 15-JAN-2002  
PF 07-AUG-2000 JP 200280989  
PR 05-AUG-1999 US 60/147499  
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EYE  
GIORDANO

```

PC C12N15/09.C12N15/09.C07K14/47.C07K16/18.C12N1/15.C12N1/19, PC
C12N1/21,
PC C12N5/10.C12P21/02.C12P21/08.C12Q1/68.C12N15/00.C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FT Key Location/Qualifiers
FT source 1..485
FT /organism='Homo sapiens (human)'
FEATURES
source 1..485
/organism="Homo sapiens"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 7,85e-68 Length: 485
Score: 74.00 Matches: 74
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.79% Indels: 0
DB: 6 Gaps: 0
US-10-023-597-24 (1-311) x BD:22339 (1-485)
QY 135 ThMetSerProGlnValCysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPhe 154
DB 263 ACCATGCTCCCGAGGTGGTGGCTCTTACGCGGTGCTACGGGATGGGGTTT 204
QY 155 GlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuVal 174
DB 203 GGGGCTGTGGCTCATACGAAATATAGTGTTCACCTTTTGCCACACACCTTGC 144
QY 175 AsnHisTyrMetCysAspIleLeuProLeuLeuGlnLeuSerCysAsnGlySerTyrIle 194
DB 143 AATCACTACATGTGTGACATCCTCCCTCTTGACCTCTCTGCAACGGCTTTACATA 84
QY 195 AsnValIleuValIlePheIleValIleThrValGlyIleGly 208
DB 63 AATGCTCTGCTCATCTTATTGTGTGACCGTTGGCATTTGGG 42
RESULT 9
AY448924 667 bp DNA linear PRI 07-DEC-2003
LOCUS Cercopithecus aethiops clone OLG_7 olfactory receptor-like protein
DEFINITION Gene, partial sequence.
ACCESSION AY448924
VERSION AY448924.1 GI:38634952
KEYWORDS
SOURCE Cercopithecus aethiops (African green monkey)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Cercopithecus.
REFERENCE 1 (bases 1 to 667)
AUTHORS Gllad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
TITLE Loss of olfactory receptor genes coupled with the acquisition of
full trichromatic vision in primates
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 667)
AUTHORS Gllad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-2003) Max Plank Institute for Evolutionary
Anthropology, Deutscher Platz 6, Leipzig 04103, Germany
LOCATION/Qualifiers
FEATURES
source 1..667
/organism="Cercopithecus aethiops"
/mol_type="genomic DNA"
/db_xref="taxon:9534"
/clone="OLG_7"
/notes="sample obtained from Primate Genetics German
Primate Center in Goettingen, Germany"
<1..>667
gene

```

```

/gene="olfactory receptor-like protein"
ORIGIN
Alignment Scores:
Pred. No.: 1,66e-63 Length: 667
Score: 70.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 22.51% Indels: 0
DB: 9 Gaps: 0
US-10-023-597-24 (1-311) x AY448924 (1-667)
QY 135 ThMetSerProGlnValCysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPhe 154
DB 215 ACCATGCTCCCGAGGTGGTGGCTCTTGTGGGTCTCATAGGAAGGAGTTT 274
QY 155 GlyAlaValAlaHisThrGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuVal 174
DB 275 GGGGCTGTGGCTCATACGAAATATAGTGTTCACCTTTTGCCACACACCTTGC 334
QY 175 AsnHisTyrMetCysAspIleLeuProLeuLeuGlnLeuSerCysAsnGlySerTyrIle 194
DB 335 AATCACTACATGTGTGACATCCTCCCTCTTGACCTCTCTGTAACGGCTTTACATA 394
QY 195 AsnValIleuValIlePheIleValIleThrValGlyIleGly 204
DB 395 AATGCTCTGCTCATCTTATTGTGTGAC 424
RESULT 10
AY448357 668 bp DNA linear PRI 07-DEC-2003
LOCUS Hylobates syndactylus clone OLG_29 olfactory receptor-like protein
DEFINITION gene, partial sequence.
ACCESSION AY448357
VERSION AY448357.1 GI:38634385
KEYWORDS
SOURCE Hylobates syndactylus (siamang)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
REFERENCE 1 (bases 1 to 668)
AUTHORS Gllad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
TITLE Loss of olfactory receptor genes coupled with the acquisition of
full trichromatic vision in primates
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 668)
AUTHORS Gllad,Y., Wiebe,V., Przeworski,M., Lancet,D. and Paabo,S.
TITLE Direct Submission
JOURNAL Submitted (20-OCT-2003) Max Plank Institute for Evolutionary
Anthropology, Deutscher Platz 6, Leipzig 04103, Germany
LOCATION/Qualifiers
FEATURES
source 1..668
/organism="Hylobates syndactylus"
/mol_type="genomic DNA"
/db_xref="taxon:9590"
/clone="OLG_29"
/notes="sample obtained from Primate Genetics German
Primate Center in Goettingen, Germany"
<1..>668
/gene="olfactory receptor-like protein"
ORIGIN
Alignment Scores:
Pred. No.: 1,74e-42 Length: 668
Score: 50.00 Matches: 50
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.08% Indels: 0
DB: 9 Gaps: 0
US-10-023-597-24 (1-311) x AY448357 (1-668)

```

QY 155 G1A1A1A1A1A1H1S1ThrG1YAsn1LeuValPheLeuThrPheCysAlaAspAsnLeuVal 174  
 DB 275 GGGGCTGGTGCATACCGGAATATAGTGTTCACCTTTGTGACAGACACCTGTC 334  
 QY 175 AsnH1s1YrMerCysAsp1LeuPheLeuGluLeuSerCysAsnGlySerTyr1Le 194  
 DB 335 AATCACTAATGTGACATCTCCCTCTTGAGCTCCTGCAACGCTCTTACATA 394  
 QY 195 AsnValLeuVal11Phe11LeuVal1Val1Thr 204  
 DB 395 AATGCTCGGTACCTTATTATTGTTGACC 424  
 RESULT 11  
 AY448759 667 bp DNA linear PRI 07-DEC-2003  
 LOCUS Cercopithecus agilis clone OLC\_56 olfactory receptor-like protein  
 DEFINITION  
 AY448759  
 ACCESSION  
 VERSION  
 KEYWORDS  
 AY448759.1 GI:38634787  
 SOURCE  
 ORGANISM  
 Cercopithecus agilis (agile mangabey)  
 Cercopithecus agilis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Cercopithecus.  
 REFERENCE  
 1 (bases 1 to 667)  
 Gilad, Y., Wiebe, V., Przeworski, M., Lancet, D. and Pabo, S.  
 Loss of olfactory receptor genes coupled with the acquisition of  
 full trichromatic vision in primates  
 Unpublished  
 JOURNAL  
 2 (bases 1 to 667)  
 Gilad, Y., Wiebe, V., Przeworski, M., Lancet, D. and Pabo, S.  
 Direct Submission  
 REFERENCE  
 Submitted (20-OCT-2003) Max Plank Institute for Evolutionary  
 Anthropology, Deutscher Platz 6, Leipzig 04103, Germany  
 JOURNAL  
 Location/Qualifiers  
 FEATURES  
 source  
 1..667  
 /organism="Cercopithecus agilis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:255237"  
 /clone="OLG\_56"  
 /note="sample obtained from Primate Genetics German  
 Primate Center in Goettingen, Germany"  
 <1..>667  
 /gene="olfactory receptor-like protein"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4.98e-33 Length: 667  
 Score: 41.00 Matches: 41  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 13.18% Indels: 0  
 DB: 9 Gaps: 0  
 US-10-023-597-24 (1-311) x AY448759 (1-667)  
 QY 164 ValPheLeuThrPheCysAlaAspAsnLeuValAsnH1s1YrMerCysAsp1LeuPro 183  
 DB 302 GGTGTTCTTACCTTTGTGAGATACCTTGTCACTACATGATGTGACATCTTCCC 361  
 QY 184 LeuLeuGluLeuSerCysAsnGlySerTyr1LeuValLeuVal11Phe11LeuVal 203  
 DB 362 CTTCCTGAGCTCTCTGTGACGCTTACATAATGCTCGTCACTTTATTGTTGTG 421  
 QY 204 Thr 204  
 DB 422 ACC 424  
 RESULT 12  
 AX646875 886 bp DNA linear PAT 04-MAR-2003  
 LOCUS  
 DEFINITION  
 Sequence 1067 from Patent EP1270724.

ACCESSION AX646875  
 VERSION AX646875.1 GI:28799398  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1  
 Suwa, M., Arai, K., Akiyama, Y. and Aburatani, H.  
 Guanosine triphosphate-binding protein coupled receptors  
 Patent: EP 1270724-A 1067 02-JAN-2003;  
 National Institute of Advanced Industrial Science and Technology  
 (NIA) ; Center for Advanced Science and Technology Incubation, Ltd.  
 (CP)  
 FEATURES  
 source  
 Location/Qualifiers  
 1..886  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 201..686  
 /note="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CAD69564.1"  
 /db\_xref="GI:28799398"  
 /db\_xref="REMBL:CAD69564"  
 /translation="MPARNSSSVTAFTLISGLTDOPGLQIPAPFLIFGYAVTVGNIG  
 LITLIGNSRLHIPIYFRPPNHSPTDFXSTLAKRMATMSVSNITISACMTQLP  
 FCFVFSSEYILSAVADRYVGINPLIIVTMSFQMLLLGLVYGMGILMLINE  
 T"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 7.18e-32 Length: 886  
 Score: 40.00 Matches: 40  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 12.86% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-023-597-24 (1-311) x AX646875 (1-886)  
 QY 28 PhePheLeuPheLeuGlyPheTyrAlaValThrVal1GlyAsnLeuGlyLeu1Le1le 47  
 DB 282 TTCCTCCCTGTTCTCTGATCCGCGGTACCGGTCAGGAGGACCTGCTGATATC 341  
 QY 48 Leu1LeuGlyLeuAsnSerTyrGluLeuH1s1LeuProMetTyrPhePheProPheAsnLeuSer 67  
 DB 342 CTGATGAGGCTCACTCTGCTGCAATCCCAATGACTTTTCCCTTCACTGTGCC 401  
 RESULT 13  
 AX241493 935 bp DNA linear PAT 26-SEP-2001  
 LOCUS  
 DEFINITION  
 Sequence 241 from Patent WO0127158.  
 AX241493  
 ACCESSION  
 VERSION AX241493.1 GI:15798368  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 synthetic construct  
 artificial sequences.  
 REFERENCE  
 1  
 Bellenson, J., Smith, D., Lancet, D., Gusman, G., Fuchs, T. and  
 Yanai, I.  
 Olfactory receptor sequences  
 Patent: WO 0127158-A 241 19-APR-2001;  
 Discents (US) ; YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)  
 Location/Qualifiers  
 1..935  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /note="(H3890 nucleotide)"  
 ORIGIN



*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 13:24:01 ; Search time 88 Seconds  
(without alignments)  
1961.247 Million cell updates/sec

Title: US-10-023-597-24

Perfect score: 311

Sequence: 1 MAENSSSTEFILAGLHQ.....LNKDVYKALKRTFSRISFS 311

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 segs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360442

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-O/cgn2.1/us3po.spool/us10023597/runat.29092004.163431.15438/app.query.fasta.1.455  
-DB=Issued\_Patents\_NA -QFMT=faststep -SUFFIX=olig.rni -MINMATCH=0.1 -LOOPCTL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORX=ext -HEAVYSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=us10023597 @CGN 1.1 69 @runat.29092004.163431.15438 -NCPU=6 -ICPU=3  
-NO NMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2.6/prodata/2/ina/5A.COMB.seq:\*  
2: /cgn2.6/prodata/2/ina/5B.COMB.seq:\*  
3: /cgn2.6/prodata/2/ina/5A.COMB.seq:\*  
4: /cgn2.6/prodata/2/ina/5B.COMB.seq:\*  
5: /cgn2.6/prodata/2/ina/PCUTS.COMB.seq:\*  
6: /cgn2.6/prodata/2/ina/backfiles1.seq:\*

Pred.No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	74	23.8	485	4	US-09-621-976-18883
C 2	13	4.2	945	4	US-09-016-434-1114
C 3	10	3.2	456	4	US-09-621-976-3067
C 4	10	3.2	561	4	US-09-621-976-3067
C 5	10	3.2	966	4	US-08-748-506-8
C 6	10	3.2	1062	4	US-09-668-80-10
C 7	10	3.2	1290	2	US-08-827-291A-1
C 8	10	3.2	1438	4	US-08-988-876-2
C 9	9	2.9	457	4	US-09-517-431E-15
C 10	9	2.9	4989	4	US-09-976-594-114
C 11	8	2.6	360	4	US-08-543-246B-4
C 12	8	2.6	405	3	US-08-543-246B-3

C 13	8	2.6	405	3	US-08-543-246B-7	Sequence 7, Appli
C 14	8	2.6	419	4	US-09-621-976-8822	Sequence 8822, Ap
C 15	8	2.6	472	4	US-09-016-434-1467	Sequence 1467, Ap
C 16	8	2.6	538	4	US-09-621-976-14167	Sequence 14167, A
C 17	8	2.6	645	3	US-08-543-246B-12	Sequence 12, Appl
C 18	8	2.6	693	3	US-08-543-246B-13	Sequence 13, Appl
C 19	8	2.6	699	3	US-08-543-246B-11	Sequence 11, Appl
C 20	8	2.6	1222	3	US-08-543-246B-5	Sequence 5, Appl
C 21	8	2.6	1223	4	US-09-016-434-1347	Sequence 1347, Ap
C 22	8	2.6	1297	4	US-09-668-680-11	Sequence 11, Appl
C 23	8	2.6	1333	3	US-08-543-246B-15	Sequence 15, Appl
C 24	8	2.6	1387	3	US-08-543-246B-1	Sequence 1, Appl
C 25	8	2.6	1406	4	US-09-149-476-13	Sequence 43, Appl
C 26	8	2.6	1755	3	US-08-543-246B-8	Sequence 8, Appl
C 27	8	2.6	1947	4	US-09-668-262A-13	Sequence 13, Appl
C 28	8	2.6	2280	4	US-09-358-352-3524	Sequence 3524, Ap
C 29	8	2.6	2688	1	US-08-088-633-3	Sequence 3, Appl
C 30	8	2.6	2688	1	US-08-245-756-3	Sequence 3, Appl
C 31	8	2.6	2688	1	US-08-441-750-3	Sequence 3, Appl
C 32	8	2.6	2688	2	US-08-441-751-3	Sequence 3, Appl
C 33	8	2.6	2688	5	PCT-US92-02521-13	Sequence 3, Appl
C 34	8	2.6	2777	3	US-09-282-147-39	Sequence 35, Appl
C 35	8	2.6	3033	4	US-09-525-160B-9	Sequence 9, Appl
C 36	8	2.6	3213	4	US-09-525-160B-4	Sequence 4, Appl
C 37	8	2.6	3459	4	US-09-016-434-1363	Sequence 1363, Ap
C 38	8	2.6	4495	4	US-09-620-312D-352	Sequence 352, Ap
C 39	8	2.6	5690	4	US-09-773-416-13	Sequence 13, Appl
C 40	8	2.6	6008	1	US-07-789-915A-5	Sequence 5, Appl
C 41	8	2.6	6008	1	US-08-005-002C-5	Sequence 5, Appl
C 42	8	2.6	6008	1	US-08-487-203A-5	Sequence 5, Appl
C 43	8	2.6	6755	3	US-08-931-999-4	Sequence 4, Appl
C 44	8	2.6	31063	4	US-09-596-002-20	Sequence 20, Appl
C 45	8	2.6	44848	4	US-09-435-739-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1				
US-09-621-976-18883/c				
Sequence 18883, Application US/09621976				
Patent No. 6639063				
GENERAL INFORMATION:				
APPLICANT: Dumas Milne Edwards, J.B.				
APPLICANT: Jobert, S.				
APPLICANT: Giordano, J.Y.				
TITLE OF INVENTION: ESTs and Encoded Human Proteins.				
FILE REFERENCE: GENSET 054PR2				
CURRENT APPLICATION NUMBER: US/09/621,976				
CURRENT FILING DATE: 2000-07-21				
NUMBER OF SEQ ID NOS: 19335				
SOFTWARE: Patent.pm				
SEQ ID NO 18883				
LENGTH: 485				
TYPE: DNA				
ORGANISM: Homo sapiens				
US-09-621-976-18883				
Alignment Scores:				
Pred. No.: 6.07e-63				
Score: 74.00				
Percent Similarity: 100.00%				
Best Local Similarity: 100.00%				
Query Match: 23.79%				
DB: 4				
Gaps: 0				
US-10-023-597-24 (1-311) x US-09-621-976-18883 (1-485)				
Cy	135	ThrtetserPrognValCysleuLeuLeuLeuGlyValTyrglYmetGlyValphe	154	
Db	263	ACCAATGCTCCCAAGTGTGTGCTCTTACTGAGTCTACGGAGATGGGGATTIT	204	
Cy	155	GlyValValAlaHisThrglYAsnIleValApheluThrPhcCysAlaAspAnLeuVal	174	



DB 203 GGGGCTGTGCTATACAGAAATATAGTGTTCACCTTTGTGACAGAACCTTGTCTC 144  
QY 175 AsnHISTYrMetCysAspIleLeuProIleuGluLeuSerCysAsnGlySerTyrIle 194  
DB 143 AATCCTACATGTGACATCTTCCTTCCTTGTAGGCTCTCTCAACGGCTCTACATA 84  
QY 195 AsnValIleuValIlePheIleValIValThrValIglyIleGly 208  
DB 83 AATGCTGCTGATCTTATGTGTGTGACCGTGGCATTGGG 42

## RESULT 2

US-09-016-434-1114  
; Sequence 1114, Application US/09016434  
; Patent No. 6500938  
; GENERAL INFORMATION:  
; APPLICANT: Janice Au-Young  
; APPLICANT: Jeffrey J. Selthamer  
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
; NUMBER OF SEQUENCES: 1490  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
; STREET: 3174 PORTER DRIVE  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/016.434  
; FILING DATE: HERewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, Karen J.  
; REGISTRATION NUMBER: 37,071  
; REFERENCE/DOCKET NUMBER: PA-0002 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 855-0555  
; TELEFAX: (650) 845-4166  
; INFORMATION FOR SEQ ID NO: 1114:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 945 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GENBANK  
; CLONE: G1336040  
; US-09-016-434-1114

## Alignment Scores:

Pred. No.: 0.00201 Length: 945  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.18% Indels: 0  
DB: 4 Gaps: 0

US-10-023-597-24 (1-311) x US-09-016-434-1114 (1-945)

QY 286 AsnProLeuIleTyrSerLeuArgAsnIysAspValIys 298  
DB 862 AATCCTGATTTATAGTTTGGAATAAAGATGTAAG 900

## RESULT 3

US-09-621-976-3066/c  
; Sequence 3066, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 3066  
; LENGTH: 456  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 171..380  
; US-09-621-976-3066

## Alignment Scores:

Pred. No.: 0.812 Length: 456  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.22% Indels: 0  
DB: 4 Gaps: 0

US-10-023-597-24 (1-311) x US-09-621-976-3066 (1-456)

QY 286 AsnProLeuIleTyrSerLeuArgAsnIys 295  
DB 424 AATCCTGATATATAGTCTCAGAACAA 395

## RESULT 4

US-09-621-976-3067/c  
; Sequence 3067, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET.054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 3067  
; LENGTH: 561  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 222..392  
; US-09-621-976-3067

## Alignment Scores:

Pred. No.: 0.993 Length: 561  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.22% Indels: 0  
DB: 4 Gaps: 0

US-10-023-597-24 (1-311) x US-09-621-976-3067 (1-561)

QY 286 AsnProLeuIleTyrSerLeuArgAsnIys 295  
DB 163 AATCCTGATATATAGTCTCAGAACAA 134

```
RESULT 5
US-08-748-506-8
; Sequence 8, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE//DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-748-506-8

Alignment Scores:
Pred. No.: 1.69 Length: 966
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.22% Indels: 0
DB: 3 Gaps: 0

US-10-023-597-24 (1-311) x US-08-748-506-8 (1-966)
QY 289 11eTySerLeuArgAsnLysAspVallys 298
Db 883 ATCTAAGTTTAGGACCAAGGATGTCAAG 912

RESULT 6
US-09-668-680-10
; Sequence 10, Application US/09668680
; Patent No. 6436703
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Xue, Aidong J.
; APPLICANT: Xu, Chongjun
; APPLICANT: Dimañac, Radoje T.
; TITLE OF INVENTION: No. 6436703e1 Nucleic Acids and
; TYPE OF INVENTION: Polypeptides
; FILE REFERENCE: 790CIP2A
```

```
; CURRENT APPLICATION NUMBER: US/09/668,680
; CURRENT FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: pl_fl_genes Version 2.0
; SEQ ID NO 10
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1062)
; US-09-668-680-10

Alignment Scores:
Pred. No.: 1.85 Length: 1062
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.22% Indels: 0
DB: 4 Gaps: 0

US-10-023-597-24 (1-311) x US-09-668-680-10 (1-1062)
QY 289 11eTySerLeuArgAsnLysAspVallys 298
Db 988 ATCTACAGCCTGAGGAACAAGATGTCAAG 1017

RESULT 7
US-08-827-291A-1
; Sequence 1, Application US/08827291A
; Patent No. 5874243
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Sathe, Ganesh
; TITLE OF INVENTION: NOVEL OLRC15 RECEPTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,291A
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T
; REGISTRATION NUMBER: 30,954
; REFERENCE//DOCKET NUMBER: GP50001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1290 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
```

MOLECULE TYPE: CDNA  
US-08-827-291A-1

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
2.23	10.00	100.00%	100.00%	3.22%	2	1290	10	0	0	0	0

US-10-023-597-24 (1-311) x US-08-827-291A-1 (1-1290)

Qy 286 AsnProLeuIleTyrSerLeuArgGlnLys 295  
Db 1151 AATCCCTCATCTACAGCTCCGCAACAAG 1180

RESULT 8

US-08-988-876-2

Sequence 2, Application US/08988876

Patent No. 6063596

GENERAL INFORMATION:

APPLICANT: Lal, Preeti

APPLICANT: Bandman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Yue, Henry

TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED

NUMBER OF SEQUENCES: 9

NUMBER OF INVENTIONS: WITH IMMUNE RESPONSE

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/988.876

FILING DATE: Herewith

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0441 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1828 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: PROSNOT01

CLONE: 364702

US-08-988-876-2

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:
3.14	10.00	100.00%	100.00%	3.22%	1828	10	0	0	0

DB: 3 Gaps: 0

US-10-023-597-24 (1-311) x US-08-988-876-2 (1-1828)

Qy 221 IleuSerSerIleuArgValSerSer 230  
Db 1447 ATTCTTCCTCATCTACAGCTCCGCAACAAG 1476

RESULT 9

US-09-517-431E-15/c

Sequence 15, Application US/09517431E

Patent No. 6593512

GENERAL INFORMATION:

APPLICANT: Vitek, Michael P.

APPLICANT: Dawson, Hana N.

APPLICANT: LORING, Jeanne F.

TITLE OF INVENTION: TRANSGENIC MOUSE EXPRESSING HUMAN TAU GENE

FILE REFERENCE: 56816-5002

CURRENT APPLICATION NUMBER: US/09/517,431E

CURRENT FILING DATE: 2000-03-02

PRIOR APPLICATION NUMBER: 60/122,691

PRIOR FILING DATE: 1999-03-02

NUMBER OF SEQ ID NOS: 22

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 15

LENGTH: 457

TYPE: DNA

ORGANISM: Homo sapiens

US-09-517-431E-15

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
7.61	9.00	100.00%	100.00%	2.89%	457	9	0	0	0	0

US-10-023-597-24 (1-311) x US-09-517-431E-15 (1-457)

Qy 99 SerGlnPhePhePhePhePhePhe 107  
Db 43 TCACAGTCTCTTTCTTTGCTTTTC 17

RESULT 10

US-09-976-594-114

Sequence 114, Application US/09976594

Patent No. 6673549

GENERAL INFORMATION:

APPLICANT: Furness, Michael

APPLICANT: Buchbinder, Jenny

TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

FILE REFERENCE: PA-0041 US

CURRENT APPLICATION NUMBER: US/09/976,594

CURRENT FILING DATE: 2001-10-12

PRIOR APPLICATION NUMBER: 60/240,409

PRIOR FILING DATE: 2000-10-12

NUMBER OF SEQ ID NOS: 1143

SOFTWARE: PERL Program

SEQ ID NO 114

LENGTH: 4989

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6673549 1382924.38

NAME/KEY: unsure

LOCATION: 99, 544

OTHER INFORMATION: a, t, c, g, or other

US-09-976-594-114

Alignment Scores:

Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	Length:	Matches:	Conservative:	Mismatches:	Indels:
78.1	10.00	100.00%	100.00%	3.22%	4989	10	0	0	0

Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.89% Indels: 0  
DB: 4 Gaps: 0

US-10-023-597-24 (1-311) x US-09-976-594-114 (1-4989)

QY 203 ValThValGlyIleGlyValProIle 211  
Db 1612 GTGACTGTAGTATCGTGTCTTCTATT 1638

RESULT 11  
US-08-543-246B-4/c  
; Sequence 4, Application US/08543246B  
; Patent No. 6262244  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: DNA and amino acid sequence specific for  
; TITLE OF INVENTION: natural killer cells  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Michael W. Glynn  
; ADDRESSEE: No. 6262244artis Corporation  
; STREET: 564 Morris Avenue  
; CITY: Summit,  
; STATE: NJ  
; COUNTRY: US  
; ZIP: 07901-1027  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/543,246B  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/676,663  
; FILING DATE: 28-MAR-1991  
; PRIOR APPLICATION DATA: PCT/US92/02469  
; APPLICATION NUMBER: 27-MAR-1992  
; FILING DATE: 24-SEP-1993  
; APPLICATION NUMBER: US 08/122,514  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kasenoff, Melvyn M.  
; REGISTRATION NUMBER: 26,389  
; REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 908-522-6927  
; TELEFAX: 908-522-6955  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 360 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; US-08-543-246B-4

Alignment Scores:  
Pred. No.: 56.4 Length: 360  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.57% Indels: 0  
DB: 3 Gaps: 0

US-10-023-597-24 (1-311) x US-08-543-246B-4 (1-360)

QY 124 ValGlyIleCysAsnProIleLeu 131  
Db 45 GTTGAAATATGTATTCATCTCTC 22

RESULT 12  
US-08-543-246B-3/c  
; Sequence 3, Application US/08543246B  
; Patent No. 6262244  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: DNA and amino acid sequence specific for  
; TITLE OF INVENTION: natural killer cells  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Michael W. Glynn  
; ADDRESSEE: No. 6262244artis Corporation  
; STREET: 564 Morris Avenue  
; CITY: Summit,  
; STATE: NJ  
; COUNTRY: US  
; ZIP: 07901-1027  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/543,246B  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/676,663  
; FILING DATE: 28-MAR-1991  
; PRIOR APPLICATION DATA: PCT/US92/02469  
; APPLICATION NUMBER: 27-MAR-1992  
; FILING DATE: 24-SEP-1993  
; APPLICATION NUMBER: US 08/122,514  
; FILING DATE: 24-SEP-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kasenoff, Melvyn M.  
; REGISTRATION NUMBER: 26,389  
; REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 908-522-6927  
; TELEFAX: 908-522-6955  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 405 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; US-08-543-246B-3

Alignment Scores:  
Pred. No.: 63.3 Length: 405  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.57% Indels: 0  
DB: 3 Gaps: 0

US-10-023-597-24 (1-311) x US-08-543-246B-3 (1-405)

QY 124 ValGlyIleCysAsnProIleLeu 131  
Db 90 GTTGAAATATGTATTCATCTCTC 67

RESULT 13  
US-08-543-246B-7/c

Sequence 7, Application US/08543246B  
Patent No. 6262244  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: DNA and amino acid sequence specific for  
TITLE OF INVENTION: natural killer cells  
NUMBER OF SEQUENCES: 24  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Michael W. Glynn  
ADDRESSEE: No. 6262244artis Corporation  
STREET: 564 Morris Avenue  
CITY: Summit,  
STATE: NJ  
COUNTRY: US  
ZIP: 07901-1027  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/543,246B  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/676,663  
FILING DATE: 28-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/02469  
FILING DATE: 27-MAR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/122,514  
FILING DATE: 24-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Kassenoff, Melvyn M.  
REGISTRATION NUMBER: 26,389  
REFERENCE/DOCKET NUMBER: 118-7704/PCT/CONT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-522-6927  
TELEFAX: 908-522-6955  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 405 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-543-246B-7  
  
Alignment Scores:  
Pred. No.: 63.3 Length: 405  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.57% Indels: 0  
DB: 3 Gaps: 0  
  
US-10-023-597-24 (1-311) x US-08-543-246B-7 (1-405)  
QY 124 ValGlyTleCyAsnProLeuLeu 131  
DB 90 GTTGAATATGTATTCACCTCCTC 67  
  
RESULT 14  
US-09-621-976-8822/C  
Sequence 8822, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Mline Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054BR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 8822  
LENGTH: 419  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-8822  
  
Alignment Scores:  
Pred. No.: 65.4 Length: 419  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.57% Indels: 0  
DB: 4 Gaps: 0  
  
US-10-023-597-24 (1-311) x US-09-621-976-8822 (1-419)  
QY 141 CysLeuLeuLeuLeuGlyVal 148  
DB 190 TGCCCTCCTGCTGCTGTGGAGTG 167  
  
RESULT 15  
US-09-016-434-1467  
Sequence 1467, Application US/09016434  
Patent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1467:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 472 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: 9902329  
US-09-016-434-1467

Alignment Scores:

Pred. No.:	73.5	Length:	472
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.57%	Indels:	0
DB:	4	Gaps:	0

US-10-023-597-24 (1-311) x US-09-016-434-1467 (1-472)

QY	223	SeSeSeIleLeuArgValSeSeSe	230
DB	293	TCCTCCATTCTGAGGCTTCAICCA	316

Search completed: September 30, 2004, 16:06:33  
Job time : 93 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 13:14:31 ; Search time 2904 Seconds  
(without alignments)  
3198.050 Million cell updates/sec

Title: US-10-023-597-24  
Perfect score: 311  
Sequence: 1 MAENSSSVTEFLAGLIHQ.....LRNDVKALAKRTSRISFS 311

Scoring table:  
OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Command line parameters:  
-MODE=frame-p2n.model -DBV=xlh  
-Q/cgn2\_1/USPTO.spool/US10023597/runat.29092004.163431.15422/app.query.fasta.1.455  
-DB=EST -QFMT=fastap -SUFFIX=01hg.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco  
-NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=300000000  
-USER=US10023597@cgn2\_1.1.3437@runat.29092004.163431.15422 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEIO -NEG\_SCORES=0 -WAIT -USPBLOCK=100 -LONGLOS  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estlu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_hiv:\*  
19: em\_gss\_pin:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_png:\*  
27: em\_gss\_vr1:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	* Query Match Length	DB ID	Description
1	21	6.8	335 14 CBI173935	CBI173935 OR 2019H05
2	21	6.8	475 28 AZ016145	AZ016145 RPT-23-3
3	20	6.4	444 14 CA861348	CA861348 K0989F04-
4	20	6.4	472 14 CA860338	CA860338 K0982C03-
5	20	6.4	799 29 AY407110	AY407110 Pan trogl
6	20	6.4	933 29 AY407109	AY407109 Homo sapi
7	20	6.4	933 29 AY407111	AY407111 Mus muscu
8	18	5.8	693 13 BY720370	BY720370 BY720370
9	16	5.1	464 14 CBI173917	CBI173917 OR 2019A0
10	16	5.1	498 28 AZ721107	AZ721107 RPT-24-1
11	16	5.1	519 14 CBI174259	CBI174259 OR 2039C1
12	16	5.1	675 28 AZ430777	AZ430777 IM0215H24
13	16	5.1	710 28 B2874853	B2874853 CH240_292
14	16	5.1	733 28 B2201201	B2201201 CH230-303
15	16	5.1	930 29 AY405659	AY405659 Mus muscu
16	15	4.8	527 14 CBI173731	CBI173731 OR 20105b
17	15	4.8	578 14 CBI174505	CBI174505 OR C10_1
18	15	4.8	637 14 CBI174340	CBI174340 OR 2042A0
19	15	4.8	642 29 CE253172	CE253172 tigr-gss-
20	15	4.8	669 28 AZ627033	AZ627033 IM0467B21
21	15	4.8	674 14 CBI174506	CBI174506 OR C11_2
22	15	4.8	771 9 A1962273	A1962273 wd46d06.x
23	15	4.8	903 29 AY405609	AY405609 Homo sapi
24	13	4.2	321 29 CC568910	CC568910 CH240_443
25	13	4.2	330 28 B2760015	B2760015 622_302.B
26	13	4.2	346 14 CBI173316	CBI173316 OR 2015E0
27	13	4.2	356 28 BH025383	BH025383 RPT-24-1
28	13	4.2	369 28 BH120437	BH120437 RPT-24-3
29	13	4.2	386 14 CBI172842	CBI172842 OR 205B11
30	13	4.2	392 14 CBI173049	CBI173049 OR 2008B0
31	13	4.2	398 14 CBI173176	CBI173176 OR E7_L10
32	13	4.2	424 29 CE056000	CE056000 tigr-gss-
33	13	4.2	432 14 CBI174364	CBI174364 OR 2041H0
34	13	4.2	454 28 AZ021177	AZ021177 RPT-23-3
35	13	4.2	458 14 CBI174053	CBI174053 OR 2032G0
36	13	4.2	471 29 CE614138	CE614138 tigr-gss-
37	13	4.2	472 14 CBI174158	CBI174158 OR 2035C0
38	13	4.2	478 28 AZ734405	AZ734405 RPT-24-1
39	13	4.2	482 29 CC472634	CC472634 CH240_296
40	13	4.2	487 28 AZ632285	AZ632285 IM0466P18
41	13	4.2	492 28 B2242329	B2242329 CH230-362
42	13	4.2	501 14 CBI174092	CBI174092 OR 2033G0
43	13	4.2	503 14 CBI173683	CBI173683 OR 2023E0
44	13	4.2	504 14 CBI174105	CBI174105 OR 2034C0
45	13	4.2	504 29 CE243689	CE243689 tigr-gss-

#### ALIGNMENTS

RESULT 1  
CBI173935  
LOCUS CBI173935 335 bp mRNA linear EST 09-OCT-2003  
DEFINITION OR 2019H05\_E02\_010105\_13.y3 Adult mouse olfactory epithelium  
1library Mus musculus cDNA clone 2019H05 5', mRNA sequence.  
ACCESSION CBI173935  
VERSION CBI173935.1 GI:37592564  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 335)

**AUTHORS**  
Young, J.M., Shykind, B.M., Lane, R.P., Tonnes-Friddy, L., Ross, J.A., Walker, M., Williams, E.M., and Trask, B.J.

**TITLE**  
Olfactory receptor expressed sequence tags demonstrate olfactory expression of over 400 genes, extensive alternate splicing and unequal expression levels

**JOURNAL**  
Genome Biol. 4 (11), R71.1-R71.15 (2003)

**COMMENT**  
Contact: Young JM  
Trask Lab, Division of Human Biology  
Fred Hutchinson Cancer Research Center  
1100 Fairview Avenue N., C3-168, P.O. Box 19024, Seattle, WA  
98109-1024, USA  
Tel: 206 667 1471  
Fax: 206 667 6524  
Email: jayoung@fhcrc.org

Young gene new name GA\_X6K02T2PVTD-31458511-31459443  
Young gene old name GA\_X5Y8B7W60AJ-1838648-1837916  
Zhang gene name MOR161-4  
Seq primer: M13 Reverse

**FEATURES**  
source

Location/Qualifiers  
1..335  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="2019H05"  
/tissue\_type="olfactory and respiratory epithelium"  
/dev\_stage="Adult"  
/clone\_lib="Adult mouse olfactory epithelium library"  
/note="Organ: Olfactory turbinates; Vector: LambdaZAPII-XR; Site 1: EcoRI; Site 2: XhoI; This library was provided by Leslie Vossahl. mRNA was prepared from the olfactory and respiratory epithelium of an adult mouse. Oligo-dT primed cDNA was directionally cloned into Stratagene's LambdaZAPII-XR vector."

#### ORIGIN

**Alignment Scores:**  
Pred. No.: 2,71e-07 Length: 335  
Score: 21.00 Matches: 21  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.75% Indels: 0  
DB: 14 Gaps: 0

US-10-023-597-24 (1-311) x CB173935 (1-335)

**QY** 259 TytleuysppropserileuProleuaspGingilysvalseSerleuphetyr 278  
Db 147 TACCTCAACCTCTTCATTTCCTGACCAAGAAAGTCTCTCTGTTTAT 206

**QY** 279 Thr 279  
Db 207 ACA 209

#### RESULT 2

AZ016145/c 475 bp DNA linear GSS 25-FEB-2000  
LOCUS RPCT-23-37915.TV RPCT-23 Mus musculus genomic clone RPCT-23-37915,  
DEFINITION genomic survey sequence.

ACCESSION  
AZ016145  
VERSION  
AZ016145.1 GI:7091529

**KEYWORDS**  
GSS.

#### SOURCE

ORGANISM  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE**  
Zhao, S., Niernan, M., Feldblum, T., Malek, J., Shatsman, S.,  
Akimov, B., Levin, M., McGinn, S., Teegsye, G., Geer, K., Krol, M., de  
Jong, P., and Frazer, C.M.

#### AUTHORS

**TITLE**  
Mouse BAC End Sequences from Library RPCT-23

**JOURNAL**  
Unpublished (1999)

**COMMENT**  
Contact: Shaying Zhao

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCT-23. For BAC library availability, please contact Pieter de Jong (pieterdejong.med.bu@falo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.bu@falo.edu/orderingframe.htm>) or from Resea.ch Genetics ([info@resgen.com](http://info@resgen.com)). BAC end page: [http://www.tigr.org/cdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html)  
Place: 379 row: 1 column: 5  
Seq primer: 17  
Class: BAC ends.

**FEATURES**  
source

Location/Qualifiers  
1..475  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCT-23-37915"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

#### ORIGIN

**Alignment Scores:**  
Pred. No.: 3.4e-07 Length: 475  
Score: 21.00 Matches: 21  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.75% Indels: 0  
DB: 28 Gaps: 0

US-10-023-597-24 (1-311) x AZ016145 (1-475)

**QY** 259 TytleuysppropserileuProleuaspGingilysvalseSerleuphetyr 278  
Db 130 TACCTCAACCTCTTCATTTCCTGACCAAGAAAGTCTCTCTGTTTAT 71

**QY** 279 Thr 279  
Db 70 ACA 68

#### RESULT 3

CA881328/c 444 bp mRNA linear EST 20-DEC-2002  
LOCUS K0989F04-SN NIA Mouse Neural Stem Cell (Undifferentiated) cDNA  
DEFINITION Library (long) Mus musculus cDNA clone NIA:K0989F04 IMAGE:30092799  
5', mRNA sequence.

ACCESSION  
CA881328  
VERSION  
CA881328.1 GI:27332877

**KEYWORDS**  
EST.

#### SOURCE

ORGANISM  
Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

**REFERENCE**  
Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Aiba, K., Vescovi, A.L.  
and Ko, M.S.H.

#### AUTHORS

**TITLE**  
Systematic Analyses of NIA Mouse Neural Stem Cell

**JOURNAL**  
Unpublished (2002)

**COMMENT**  
Laboratory of Genetics

National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
Plate: K0989 row: F column: 04  
Seq primer: -21M13 Reverse  
High quality sequence stop: 444  
POLYA=NO.

## FEATURES

source

Location/Qualifiers  
1..444  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CD1"  
/db\_xref="niaEST:K0989F04-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0989F04 IMAGE:30092799"  
/dev\_stage="Adult"  
/lab\_host="DHI0B"  
/clone\_lib="NIA Mouse Neural Stem Cell (Undifferentiated)  
CDNA Library (Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11:1553-1558 (2001). [PMID: 11544199]). Total RNAs were  
obtained from Dr. Angelo L. Vescovi (Institute for Stem  
Cell Research, Italy). Double-stranded cDNAs were  
synthesized with an oligo(dT) primer [Invitrogen:  
5'-pGACTGATCTAGATCGAGCGCGCCCTTTTCTTTT-3'] from  
2.0 Microgram of total RNA, treated with T4 DNA  
polymerase, and purified by ethanol-precipitation. The  
cDNAs were ligated to lone-linker L1-Sal4, purified by  
phenol/chloroform, and separated from free linkers by  
Centricon 100. Then, the cDNAs were amplified by  
long-range high fidelity PCR using Ex Taq polymerase  
(Takara) with a primer Sal4-S. The products were purified  
by phenol/chloroform and Centricon 100. The cDNAs were  
digested with SalI and NotI enzymes and cloned into  
SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.  
coli host was transformed with the ligation mixture by the  
standard chemical method. The average insert size is about  
3.8 kb. The library was constructed by Yulan Piao."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,13e-06 Length: 444  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 14 Gaps: 0

US-10-023-597-24 (1-311) x CA881328 (1-444)

QY 259 TyrlenuysProProserilleuProluaspGlnGlyValserseuPheTyr 278

DB 157 TACCTCAAACTCTTCATTTGCTCTTGACCAAGAAAGTCTTCTGTTCTAT 98

## RESULT 4

CA880338 472 bp mRNA linear EST 20-DEC-2002  
LOCUS CA880338  
DEFINITION K0982C03-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA  
Library (Long) Mus musculus cDNA clone NIA:K0982C03 IMAGE:30092090  
5', mRNA sequence.

ACCESSION CA880338.1 GI:27331887  
VERSION CA880338  
KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 472)  
REFERENCE  
AUTHORS Piao,Y., Dudekula,D.B., Qian,Y., Martin,P.R., Alba,K., Vescovi,A.L.

TITLE and Ko,Y.S.H.  
JOURNAL Systematic Analyses of NIA Mouse Neural Stem Cell  
COMMENT (Undifferentiated) cDNA Library (Long)  
Unpublished (2002)  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@igsun.grc.nia.nih.gov  
Plate: K0982 row: C column: 03  
Seq primer: -21M13 Reverse  
High quality sequence stop: 472  
POLYA=NO.

## FEATURES

source

Location/Qualifiers  
1..472  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="CD1"  
/db\_xref="niaEST:K0982C03-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0982C03 IMAGE:30092090"  
/dev\_stage="Adult"  
/lab\_host="DHI0B"  
/clone\_lib="NIA Mouse Neural Stem Cell (Undifferentiated)  
CDNA Library (Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site\_1: SalI; Site\_2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11:1553-1558 (2001). [PMID: 11544199]). Total RNAs were  
obtained from Dr. Angelo L. Vescovi (Institute for Stem  
Cell Research, Italy). Double-stranded cDNAs were  
synthesized with an oligo(dT) primer [Invitrogen:  
5'-pGACTGATCTAGATCGAGCGCGCCCTTTTCTTTT-3'] from  
2.0 Microgram of total RNA, treated with T4 DNA  
polymerase, and purified by ethanol-precipitation. The  
cDNAs were ligated to lone-linker L1-Sal4, purified by  
phenol/chloroform, and separated from free linkers by  
Centricon 100. Then, the cDNAs were amplified by  
long-range high fidelity PCR using Ex Taq polymerase  
(Takara) with a primer Sal4-S. The products were purified  
by phenol/chloroform and Centricon 100. The cDNAs were  
digested with SalI and NotI enzymes and cloned into  
SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.  
coli host was transformed with the ligation mixture by the  
standard chemical method. The average insert size is about  
3.8 kb. The library was constructed by Yulan Piao."

## ORIGIN

## Alignment Scores:

Pred. No.: 2,22e-06 Length: 472  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 14 Gaps: 0

US-10-023-597-24 (1-311) x CA880338 (1-472)

QY 259 TyrlenuysProProserilleuProluaspGlnGlyValserseuPheTyr 278

DB 157 TACCTCAAACTCTTCATTTGCTCTTGACCAAGAAAGTCTTCTGTTCTAT 98

## RESULT 5

AY407110 799 bp DNA linear GSS 15-DEC-2003  
LOCUS AY407110  
DEFINITION Pan troglodytes ORB12 gene, VIRUAL TRANSCRIPT, partial sequence.  
genomic survey sequence.

ACCESSION AY407110  
VERSION AY407110.1 GI:39763081  
KEYWORDS GSS.

## SOURCE

Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes  
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
REFERENCE 1 (bases 1 to 799)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 799)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..799  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>799  
/gene="OR8B12"  
/locus\_tag="HCM2781"  
gene  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,12e-06 Length: 799  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 29 Gaps: 0  
US-10-023-597-24 (1-311) x AY407110 (1-799)  
Qy 79 ProlysmelluMetSerPheValSerArglyAsnIleIleSerPheThgIcyMet 98  
Db 232 CCCAAATGCTGATGAGTTTGTCTCCAGAGAAACATCTTCCTTCACAGGGGTGATG 291  
RESULT 6  
LOCUS AY407109 933 bp DNA linear GSS 15-DEC-2003  
DEFINITION Homo sapiens OR8B12 gene, VIRUAL TRANSCRIPT, partial sequence,  
KEYWORDS genomic survey sequence.  
ACCESSION AY407109  
VERSION AY407109.1 GI:39763080  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 933)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 933)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment  
FEATURES  
source Location/Qualifiers  
1..933  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>933  
/gene="OR8B12"  
/locus\_tag="HCM2781"  
gene  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,45e-06 Length: 933  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 29 Gaps: 0  
US-10-023-597-24 (1-311) x AY407109 (1-933)  
Qy 79 ProlysmelluMetSerPheValSerArglyAsnIleIleSerPheThgIcyMet 98  
Db 232 CCCAAATGCTGATGAGTTTGTCTCCAGAGAAACATCTTCCTTCACAGGGGTGATG 291  
RESULT 7  
LOCUS AY407111 933 bp DNA linear GSS 15-DEC-2003  
DEFINITION Mus musculus OR8B12 gene, VIRUAL TRANSCRIPT, partial sequence,  
KEYWORDS genomic survey sequence.  
ACCESSION AY407111  
VERSION AY407111.1 GI:39763082  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 933)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE 2 (bases 1 to 933)  
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
TITLE Direct Submission  
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..933  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..>933  
/gene="OR8B12"  
/locus\_tag="HCM2781"  
gene  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,45e-06 Length: 933  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multiplexillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.

FEATURES  
    source  
        Location/Qualifiers  
            1..693  
                /organism="Mus musculus"  
                /mol\_type="mRNA"  
                /strain="C57BL/6J"  
                /db\_xref="taxon:10090"  
                /clone="7420497B16"  
                /sex="Female"  
                /tissue\_type="in vitro fertilized eggs"  
                /dev\_stage="egg"  
                /lab\_host="DH103"  
                /clone\_1pb="RIKEN full-length enriched, in vitro  
                    fertilized eggs"  
                /note="Site 1: Salt; Site 2: BamHI; cDNA library was  
                    prepared and sequenced in Mouse Genome Encyclopedia  
                    Project of Genome Exploration Research Group in Riken  
                    Genomic Sciences Center and Genome Science Laboratory in  
                    RIKEN. Division of Experimental Animal Research in Riken  
                    contributed to prepare mouse tissues. 1st strand cDNA was  
                    primed with a primer [5'  
                    GAGAGAGAGATCTCCGATTCACAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was  
                    prepared by using trehalose thermo-activated reverse  
                    transcriptase and subsequently enriched for full-length by  
                    cap-trapper. Second strand cDNA was prepared with the  
                    primer adapter of sequence [5'  
                    GAGAGAGATCTCTCGATTAATTAATTAATCCCCCCCCCCC 3']. cDNA  
                    was cloned into the XhoI and BamHI sites. Vector: a  
                    modified pluescript KS(+) after bulk excision from Lambda  
                    FLC I. Cloning sites, 5' end: Salt; 3' end: BamHI"

ORIGIN

Alignment Scores:

Pred. No.:	0.000122	Length:	693
Score:	18.00	Matches:	18
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.79%	Indels:	0
DB:	13	Gaps:	0

US-10-023-597-24 (1-311) x BY720370 (1-693)

QY 281 ValValaProMetPheasProLeuIleItyrSerLeuArganlysaAPVallys 298  
|||||  
DB 2 GTGGTGGCCCATGTTCATTCATTATATATAGCCTGAGCATATAGGATGTCAAA 55  
|||||

RESULT 9  
LOCUS CB173917 464 bp mRNA linear EST 09-OCT-2003  
DEFINITION OR\_2019A03\_E01\_010105\_05\_Y3 Adult mouse olfactory epithelium  
library Mus musculus cDNA clone 2019A03 5', mRNA sequence.  
ACCESSION CB173917  
VERSION CB173917.1 GI:37592546  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus (house mouse)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 464)  
Young, J.M., Shykind, B.M., Lane, R.P., Tonnes-Priddy, L., Ross, J.A.,  
Walker, M., Williams, E.M. and Trask, B.J.  
Odorant receptor expressed sequence tags demonstrate olfactory  
expression of over 400 genes, extensive alternate splicing and  
unequal expression levels  
Genome Biol. 4 (11), R71.1-R71.15 (2003)  
Contact: Young JM  
Trask Lab, Division of Human Biology  
Fred Hutchinson Cancer Research Center  
1100 Fairview Avenue N., C3-168, P.O. Box 19024, Seattle, WA  
98109-1024, USA  
Tel: 206 667 1471  
Fax: 206 667 6524  
Email: jayoung@fhcrc.org  
Young gene new name GA\_x6K0212PVTD-32518237-32519172  
Young gene old name GA\_x508B7W60AJ-78979-788844  
Zheng gene name MOR161-3  
Seq primer: M13 Reverse.

FEATURES  
source

Location/Qualifiers  
1..464  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="2019A03"  
/tissue\_type="olfactory and respiratory epithelium"  
/dev\_stage="Adult"  
/clone\_lib="Adult mouse olfactory epithelium library"  
/note="Organ: Olfactory tubinates; Vector:  
lambdaZAP1-XR; Site\_1: EcoRI; Site\_2: XhoI. This library  
was provided by Leslie Vossahl. mRNA was prepared from  
the olfactory and respiratory epithelium of an adult  
mouse. Oligo-dT primed cDNA was directionally cloned into  
Stratagene's lambdaZAP1-XR vector."

## ORIGIN

## Alignment Scores:

Pred. No.: 0.00402 Length: 464  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.14% Indels: 0  
DB: 14 Gaps: 0

US-10-023-597-24 (1-311) x CB173917 (1-464)

Qy 174 ValaSNH:STYMETCYASPIIleuPProleuEngiuleuSerCys 189  
Db 112 GTCAATCATTCACATGTGACATCTCTCTGCTAGAGCTTCTCTGC 159

RESULT 10  
A2721107 498 bp DNA linear GSS 24-JAN-2001  
LOCUS RPCI-24-156E15.TV RPCI-24 Mus musculus genomic clone  
DEFINITION RPCI-24-156E15, genomic survey sequence.  
ACCESSION A2721107  
VERSION A2721107.1 GI:12463469  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 498)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Zhao, S., Nietman, W., Malek, J., Shatsman, S., Akintc, B., Levins, M.,  
Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorjls, E.,  
Russell, D., de Jong, P. and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-24  
Unpublished (1999)  
Other GSSs: RPCI-24-156E15.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research

FEATURES  
source

9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0206  
Fax: 301 838 0208  
Email: szhou@igr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@mail.cho.org). Clones may be purchased from BACPAC  
Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end  
page: [http://www.tigr.org/cdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/cdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 156 row: E column: 15  
Seq primer: 17  
Class: BAC ends.  
Location/Qualifiers  
1..498  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-156E15"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/clone\_lib="RPCI-24"  
/note="Vector: pTRABAC1, Site\_1: BamHI, Site\_2: BamHI,  
RPCI-24 Mouse BAC library produced by Pieter de Jong. The  
library was cloned in the pTRABAC1 cloning vector at the  
BamHI sites using MboI partially digested male C57BL/6J  
DNA."

## ORIGIN

## Alignment Scores:

Pred. No.: 0.00421 Length: 498  
Score: 16.00 Matches: 16  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.14% Indels: 0  
DB: 28 Gaps: 0

US-10-023-597-24 (1-311) x A2721107 (1-498)

Qy 37 ValThVaIvaIaGlyASnleuGlyleuIleleuIleGlyleuasn 52  
Db 338 GTCAACAGTGTGTGGAACCTGGCGTTGATCATCCGATTGGCTCAAT 291

RESULT 11  
CB174259 519 bp mRNA linear EST 09-OCT-2003  
LOCUS OR\_2039C11\_011004.y1 Adult mouse olfactory epithelium library Mus  
DEFINITION musculus cDNA clone 2039C11 5', mRNA sequence.  
ACCESSION CB174259  
VERSION CB174259.1 GI:37592886  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 519)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Young, J.M., Shykind, B.M., Lane, R.P., Tonnes-Priddy, L., Ross, J.A.,  
Walker, M., Williams, E.M. and Trask, B.J.  
Odorant receptor expressed sequence tags demonstrate olfactory  
expression of over 400 genes, extensive alternate splicing and  
unequal expression levels  
Genome Biol. 4 (11), R71.1-R71.15 (2003)  
Contact: Young JM  
Trask Lab, Division of Human Biology  
Fred Hutchinson Cancer Research Center  
1100 Fairview Avenue N., C3-168, P.O. Box 19024, Seattle, WA  
98109-1024, USA  
Tel: 206 667 1471  
Fax: 206 667 6524  
Email: jayoung@fhcrc.org  
Young gene new name GA\_x6K02T20125-47904535-47903615  
Young gene old name GA\_x508B7W106G-2523-3443  
Young gene old name GA\_x508B7W4P0Y-571-1491

Zhang Gene name MOR150-3P  
Seq primer: M13 Reverse.

## FEATURES

Location/Qualifiers  
1. 519  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="2039c11"  
/tissue\_type="olfactory and respiratory epithelium"  
/dev\_stage="Adult"  
/note="Organ: Olfactory turbinates; Vector: LambdaZAPII-XR; Site 1: EcoRI; Site 2: XhoI; This library was provided by Leslie Vossahl. mRNA was prepared from the olfactory and respiratory epithelium of an adult mouse. Oligo-dT primed cDNA was directionally cloned into Stratagene's LambdaZAPII-XR vector."

## ORIGIN

## Alignment Scores:

Pred. No.:	0.00432	Length:	519
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.14%	Indels:	0
DB:	14	Gaps:	0

US-10-023-597-24 (1-311) x CB174259 (1-519)

QY 283 PrcMetPheAsnProLeuIleTySerLeuArgAsnLyaspVallys 298

DB 454 CCCATGTTTATCCCTGATCTACAGCTTGAGAAATTAAGATGTAATA 501

## RESULT 12

AZ430777 675 bp DNA linear GSS 03-OCT-2000  
LOCUS AZ430777.1  
DEFINITION 1M0215H24F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
clone UUGC1M0215H24 F, genomic survey sequence.

ACCESSION AZ430777  
VERSION AZ430777.1  
KEYWORDS GSS.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 675)  
Dunn,D., Aoyagi,A., Barber,M., Beacom,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weis,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Seq primer: CGTGTGTAACGACGCGCGCT  
Class: plasmid ends  
High quality sequence stop: 675.  
Location/Qualifiers  
1. 675  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0215H24"

## FEATURES

Location/Qualifiers  
1. 675  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0215H24"

## ORIGIN

## Alignment Scores:

Pred. No.:	0.00513	Length:	675
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.14%	Indels:	0
DB:	28	Gaps:	0

US-10-023-597-24 (1-311) x AZ430777 (1-675)

QY 286 AsnProLeuIleTySerLeuArgAsnLyaspVallysLeuAlaLeu 301

DB 326 AACCTTATATCTACGTTTGAGAAATTAAGATGTTAACTTGCCCTG 279

## RESULT 13

BZ874853 710 bp DNA linear GSS 18-MAR-2003  
LOCUS BZ874853.1  
DEFINITION CH240.292D3.TU CHORI-240 Bos taurus genomic clone CH240\_292D3,  
genomic survey sequence.

ACCESSION BZ874853  
VERSION BZ874853.1  
KEYWORDS GSS.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus

REFERENCE  
AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos. 1 (bases 1 to 710)  
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn,A., Gebregorgis,E., Chen,D., Riggs,F., de Jong,P., Crawford,A.M. and McEwan,J.C.

TITLE Bovine BAC End Sequences from Library CHORI-240  
JOURNAL Unpublished (2003)  
COMMENT Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208

Email: szhao@igf.org  
Clones are derived from the bovine BAC library CHORI-240  
(http://www.chori.org/bacpac/bovine240.htm). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/ordering/information.htm). This work  
was undertaken as part of the International Bovine BAC Mapping  
Consortium (IBMC) by AgResearch Ltd., New Zealand and The  
Institute of Genomic Research (TIGR), USA.  
Plate: 292 row: D column: 3



Seq primer: SP6  
Class: BAC ends.  
FEATURES  
source  
location/Qualifiers

1..710  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/strain="bred: Hereford"  
/db\_xref="taxon:9913"  
/clone="CH240\_292D3"  
/sex="Male"  
/cell\_type="Blood"  
/clone\_1ib="CHORI-240"  
/note="Vector: pTARBAcl.3; Site 1: MboI; Site 2: MboI; Hereford bull 11 Domino 99375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

## ORIGIN

## Alignment Scores:

Pred. No.:	0.0053	Length:	710
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.14%	Indels:	0
DB:	28	Gaps:	0

US-10-023-597-24 (1-311) x BZ874853 (1-710)

Qy 286 AsnProLeuIleTySerLeuArgAsnLysAspValLysLeuAlaLeu 301

Db 616 AATCCTTATTATTAAGCTTAGGACAAAGATGTTAAGCTTGCTG 569

RESULT 14  
BZ201201/c BZ201201 733 bp DNA linear GSS 11-OCT-2002

LOCUS CH230-303F3.TV CHORI-230 Segment 2 Rattus norvegicus genomic clone  
DEFINITION CH230-303F3, genomic survey sequence.

ACCESSION BZ201201

VERSION BZ201201.1 GI:23859253

KEYWORDS GSS.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 733)

AUTHORS Zhao,S., Shetty,J., Shattman,S., Tsegaye,G., Geer,K.,  
Shvartbeyn,A., Gebregeorgis,B., Overton,L., Russell,D., Chen,D.,  
Riggs,F., de Jong,P. and Fraser,C.M.

TITLE Rat BAC End Sequences from Library CHORI-230 MboI segment

JOURNAL Unpublished (1999)

COMMENT Other GSSs: CH230-303F3.TU

CONTACT: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@igf.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pdejong@mai.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or ordering information.htm). BAC end

page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html

Plate: 303 row: F column: 3

Seq primer: T7

Classes: BAC ends.

FEATURES  
source  
location/Qualifiers

1..733  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SsNHsd/MCM"  
/db\_xref="taxon:10116"

/clone="CH230-303F3"

/sex="Female"

/cell\_type="Brain"

/clone\_1ib="CHORI-230 Segment 2"

/note="Vector: pTARBAcl.3; Site 1: MboI; Site 2: MboI;  
CHORI-230 Rat (BN/SsNHsd/MCM) BAC library produced by  
Pieter de Jong"

## ORIGIN

## Alignment Scores:

Pred. No.:	0.00541	Length:	733
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.14%	Indels:	0
DB:	28	Gaps:	0

US-10-023-597-24 (1-311) x BZ201201 (1-733)

Qy 286 AsnProLeuIleTySerLeuArgAsnLysAspValLysLeuAlaLeu 301

Db 93 AATCCTTATTATTAAGCTTAGGACAAAGATGTTAAGCTTGCTTG 46

RESULT 15  
AY405659

LOCUS AY405659 930 bp DNA linear GSS 15-DEC-2003

DEFINITION Mus musculus OR884 gene, VIRUAL TRANSCRIPT, partial sequence,

ACCESSION AY405659

VERSION AY405659.1 GI:39761633

KEYWORDS GSS.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 930)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tannenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 930)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
Todd,M.A., Tannenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering

them based on alignment.

FEATURES  
source  
location/Qualifiers

1..930

/organism="Mus musculus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10090"

/gene="OR884"

/locus\_tag="HMCW2297"

## ORIGIN

## Alignment Scores:

Pred. No.:	0.00631	Length:	930
Score:	16.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.14%	Indels:	0
DB:	29	Gaps:	0

US-10-023-597-24 (1-311) x AY405659 (1-930)

us-10-023-597-24.01g.rst

Page 9

286 AsnProLeuIleTYrSerLeuAIRAsnLysAspValLysLeuIleu 301  
853 AACCACTGATCTATAGTTGAGCAATAAGACGTTAAACTTGCCTA 900

Search completed: September 30, 2004, 16:04:59  
Job time : 2910 secs

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: September 30, 2004, 12:11:46; Search time 447 Seconds  
(without alignments)  
2955.681 Million cell updates/sec

Title: US-10-023-597-24  
Perfect score: 311  
Sequence: 1 MAENSSVTEFLAGLIHQ.....LRNDVKALKRTFSRISFS 311

Scoring table:  
OLIGO -  
Xgapop 60.0, Xgapext 60.0  
Ygapop 60.0, Ygapext 60.0  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745144

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USFTO.spool/US10023597/runat.29092004\_163430\_15400/app.query.fasta\_1.455  
-DB=N.Geneseq.29Jan04 -QFWT=fastcd -SUFFIX=olig.rng -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=us10023597 @CGN 1.1 470 -runat.29092004\_163430\_15400 -NCPU=6 -ICPU=3  
-NO MAP LARGEOUTERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

1: Geneseq.29Jan04:\*  
2: geneseqn1980s:\*  
3: geneseqn1990s:\*  
4: geneseqn2000s:\*  
5: geneseqn2001as:\*  
6: geneseqn2001bs:\*  
7: geneseqn2002as:\*  
8: geneseqn2003as:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	311	100.0	963	6 AAF88413	AAf88413 Human GPC
2	210	67.5	994	6 AAH31675	AAh31675 Human Olf
3	191	61.4	931	9 ADC86616	Adc86616 Human GPC
4	156	50.2	960	9 ADD12748	Adi12748 CDNA enco
5	40	12.9	886	9 ADC86614	Adc86614 Human GPC
6	40	12.9	935	4 AAH31668	AAh31668 Human Olf
7	33	10.6	487	4 AAH84099	AAh84099 Mouse Olf
8	20	6.4	485	4 AAH84021	AAh84021 Eulemar F

9	20	6.4	930	4 AAH32421	AAh32421 Human Olf
10	20	6.4	933	5 AAS15910	AAs15910 DNA encod
11	20	6.4	933	5 AAS42233	AAs42233 Human CDN
12	20	6.4	933	6 ABZ43067	ABz43067 Human GPC
13	20	6.4	933	6 ABK68455	ABk68455 Human DNA
14	20	6.4	933	6 ABK37521	ABk37521 DNA encod
15	20	6.4	953	6 ABT05655	ABt05655 GPCR 4 pr
16	20	6.4	976	6 ABK97207	ABk97207 Human G-P
17	20	6.4	976	9 AAD60328	AAd60328 Human G-P
18	20	6.4	991	6 ABSS58782	ABs58782 Human G-P
19	20	6.4	1333	9 ADC85838	ADc85838 Human GPC
20	20	6.4	1655	6 AAD24449	AAd24449 Human G-P
21	19	6.1	486	4 AAH84095	AAh84095 Mouse Olf
22	19	6.1	487	4 AAH84089	AAh84089 Saimirib
23	19	6.1	988	6 AAF88435	AAf88435 Human GPC
24	19	6.1	958	9 ADD12792	Adi12792 CDNA enco
25	19	6.1	993	6 ABQ88089	ABq88089 Human GPC
26	19	6.1	993	6 ADC79407	Adc79407 Human G P
27	18	5.8	487	4 AAH84062	AAh84062 Pongo pyg
28	18	5.8	1015	6 ABQ88091	ABq88091 Human GPC
29	17	5.5	486	4 AAH84022	AAh84022 Eulemar F
30	17	5.5	489	4 AAH84050	AAh84050 Macaca sy
31	17	5.5	871	9 ADC86676	Adc86676 Human GPC
32	17	5.5	937	6 ABS58846	ABs58846 Human G-P
33	17	5.5	937	6 ABQ88360	ABq88360 Human G P
34	17	5.5	973	4 AAH32263	AAh32263 Human Olf
35	16	5.1	934	6 AAF88464	AAf88464 Human GPC
36	16	5.1	934	9 ADD12850	Adi12850 CDNA enco
37	16	5.1	953	6 AAF88433	AAf88433 Human GPC
38	16	5.1	1014	9 ADC79411	Adc79411 Human G P
39	15	4.8	798	4 AAH31604	AAh31604 Human Olf
40	15	4.8	801	6 ABZ43085	ABz43085 Human GPC
41	15	4.8	933	4 AAH32246	AAh32246 Human Olf
42	15	4.8	933	6 ABK68672	ABk68672 Human DNA
43	15	4.8	936	5 AAS15911	AAs15911 DNA encod
44	15	4.8	936	6 ABK40195	ABk40195 Human G P
45	15	4.8	942	6 ABZ43068	ABz43068 Human GPC

#### ALIGNMENTS

RESULT 1	AAf88413	standard; CDNA; 963 BP.
XX	AAf88413	
XX	AAf88413	
DT	12-NOV-2002	(first entry)
DE	Human GPCRX CDNA SEQ ID 23.	
XX		
XX		
XX		
OS	Homo sapiens.	
XX		
PN	W0200250275-A2.	
XX		
PD	27-JUN-2002.	
XX		
PF	18-DEC-2001; 2001WO-US048958.	
XX		
PR	18-DEC-2000; 2000US-0256635P.	
PR	21-DEC-2000; 2000US-0257876P.	
PR	04-JAN-2001; 2001US-0259743P.	

PR 10-JAN-2001; 2001US-0260718P.  
 PR 12-JAN-2001; 2001US-0264498P.  
 PR 24-JAN-2001; 2001US-0265689P.  
 PR 08-FEB-2001; 2001US-0267464P.  
 PR 22-FEB-2001; 2001US-0271021P.  
 PR 14-MAR-2001; 2001US-0275946P.  
 PR 23-MAR-2001; 2001US-0278150P.  
 PR 18-APR-2001; 2001US-0284591P.  
 PR 23-APR-2001; 2001US-0285718P.  
 PR 19-JUN-2001; 2001US-0299327P.  
 PR 16-AUG-2001; 2001US-0312902P.  
 (CUBA-) CUBAGEN CORP.

XX Padigaru M, Kekuda R, Li L, Ballinger RA, Casman SU, Spytek KA;  
 PI Baumgartner JC, Burgess CE;  
 XX MPI, 2002-657419/70.  
 DR P-PSDB; AAB71174.

PT New G-protein coupled receptor polypeptides, useful in gene therapy,  
 PT particularly for treating or preventing cardiomyopathy, atherosclerosis,  
 PT diabetes, Crohn's disease, hemophilia or cancer in humans.

XX Claim 3; Page 90; 155pp; English.

XX This invention describes novel human G-protein coupled receptor (GPCR)  
 CC polypeptides which have anti-HIV, nootropic, antiallergic, protozoal,  
 CC antiartherosclerotic, immunosuppressive, immunomodulatory, cytostatic,  
 CC antiinflammatory, antidiabetic, neuroprotective, anorectic, haemostatic,  
 CC antibacterial, fungicidal and vincidic activity. The products of the  
 CC invention can be used in gene therapy or for vaccines. The GPCR  
 CC polypeptide, GPCR nucleic acid and antibody are useful for treating,  
 CC preventing or alleviating a GPCR-associated disorder or a pathological  
 CC state in a subject e.g. cardiomyopathy, atherosclerosis, diabetes, or a  
 CC disorder related to cell signal processing and metabolic pathway  
 CC modulation. The GPCR polypeptide and nucleic acid are also useful for  
 CC diagnosing the presence of or predisposition to a disease associated with  
 CC altered levels of GPCR, particularly cancer. These polypeptides, nucleic  
 CC acids and antibodies are also useful for treating or preventing obesity,  
 CC neurodegenerative disorders, AIDS, cancer-associated cachexia, anorexia,  
 CC Alzheimer's disease, Parkinson's disease, immune disorders, graft versus  
 CC host disease, bronchial asthma, Crohn's disease, multiple sclerosis,  
 CC haemophilia, idiopathic thrombocytopenic purpura or infectious diseases.  
 CC They can also be used to screen for potential agonist and antagonist  
 CC compounds. The polypeptides are also useful as immunogens to produce  
 CC antibodies or as vaccines. Anti-GPCR antibodies can be used  
 CC diagnostically to monitor protein levels in tissue as part of a clinical  
 CC testing procedure such as in determining the efficacy of a given  
 CC treatment regimen. The host cells are useful in producing non-human  
 CC transgenic animals which are useful for studying the function and/or  
 CC activity of GPCR protein and for identifying and/or evaluating  
 CC modulators of GPCR protein activity. AAF8462-AAF8465 encode the human  
 CC GPCR proteins represented in AAB71163-AAB71226

XX Sequence 963 BP; 193 A; 247 C; 204 G; 319 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	341e-288	Length:	963
Score:	311.00	Matches:	311
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

US-10-023-597-24 (1-311) X AAF88413 (1-963)

QY 1 MetAlaAlaGluSerSerSerValThrGluPheIleLeuAlaGlyLeuIleHisGln 20  
 DB 2 ATGGCTGGCGGAATCTCTCTCCCTGACAGATTATCTCGGAGGCTTATATCCACAG 61  
 QY 21 ProGlyLeuGlnValProValPhePheLeuPheLeuGlyPheTyrAlaValThrValVal 40

DB 62 CCGGACTCCAGTCCCCCTCTTCTCTGTTCTAGGTTTCTACGGGTACGGGTG 121  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
 DB 122 GGGAACTGGGCTTGATTAATCTGATAGGCTCAACTCGCTGCATATCCCATGATAC 181  
 QY 61 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrIleIleProlys 80  
 DB 182 TTTTTCCTTCAACTGTCCTCCGTCAGATTATGTTCTCTACCAACATCATTTCCCAA 241  
 QY 81 MetLeuMetSerPheValSerArgPheAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 242 ATGCTAGTAGTTTGTCTCAAGGAAGAACATATTATTTCTTACAGGGGTGATGATCAG 301  
 QY 101 PhePhePhePheCysPhePheValPheSerGlnSerPheIleLeuSerAlaMetValGln 120  
 DB 302 TTCTTCTTCTGTTCTTGTCTTGTCTTGTGAGTCTTCATCTGTCGCGATGGTGAG 361  
 QY 121 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIleThrMetSerProGlnVal 140  
 DB 362 GACCGCTAGTGGGATCTGTAACCCACTGTGTACAGATCACCATGCTCCCAAGTG 421  
 QY 141 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
 DB 422 TGTTCCTCTTTTACTGGGTGTACGGGATGGGGTTTGGGGCTGGCTCATACA 481  
 QY 161 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
 DB 482 GGAATATATGTTTCTCTACCTTTGTGCAGCAACCTTGCAACCTGATCATATGTGTAC 541  
 QY 181 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
 DB 542 ATCTTCCCTCTTGTAGCTCTCCGCAAGGCTTTCAATGATGCTGATCTTT 601  
 QY 201 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
 DB 602 ATGTGTGTACCGTTGGGATGGGGTGGCCATTTGTCCGTTTATCTTATGCTTAT 661  
 QY 221 IleLeuSerSerIleLeuLeuValSerSerArgIleGlyIleArgSerValAlaPheSerSer 240  
 DB 662 ATCTTTCAGCATTTCTCGGCTTGTCTGTAGGGCAGGCTTAAAGCTTTCAGTAC 721  
 QY 241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260  
 DB 722 TCCAGCTCATATATGACAGTTCTCTTCTTGGGTACAGGAGCTTTTACGTAACCTC 781  
 QY 261 IysProProSerIleLeuProLeuAspGlnGlyIleValSerSerLeuPheTyrThrThr 280  
 DB 782 AAACCCCTTCATTTTACCTTGACCAAGGGAAGTGTCTCCCTGTCTATACACT 841  
 QY 281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAla 300  
 DB 842 GTGGTCCCATGTTTAACTCATTAATCTACAGCTCGAAGGAATTAAGAGTCAAACTGCC 901  
 QY 301 LeuLysArgThrPheSerArgIleSerPheSer 311  
 DB 902 CTGAAGAGAACCTTTTCCAGATTAAGCTTTTCT 934

RESULT 2  
 AAH31675  
 ID AAH31675 standard; DNA; 994 BP.  
 AC AAH31675;  
 XX 30-JUL-2001 (first entry)  
 DE Human olfactory receptor polynucleotide, SEQ ID NO: 248.  
 XX Human olfactory receptor; OR; primary scent determination;  
 KW secondary scent determination; polypeptide library; odour receptor;  
 KW scent profile; scent fingerprint; scent representation; ds.  
 XX Homo sapiens.

XX WO200127158-A2.  
 XX 19-APR-2001.  
 PD 06-OCT-2000; 2000WO-US027582.  
 XX 08-OCT-1999; 99US-015861SP.  
 PR 24-FEB-2000; 2000US-0184809P.  
 XX (DIGI-) DIGISCENTS.  
 PA (YEDA ) YEDA RES & DEV CO LTD.  
 PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
 XX WPI; 2001-290713/30.  
 DR New polynucleotides which encode polypeptides involved in olfactory  
 PT sensation for identifying olfactory agonists and antagonists.  
 XX  
 PS Claim 8; Page 276; 1857pp; English.  
 CC The present sequence is one of a number of isolated polynucleotides which  
 CC encode polypeptides involved in olfactory sensation. The polynucleotides  
 CC can be used in screening for olfactory agonists and antagonists. The  
 CC methods allow for the determination of primary scents and the  
 CC identification of the odour receptors used to detect these primary  
 CC scents. The methods also enable determination of secondary scents and the  
 CC identification of combinations of odour receptors that are involved in  
 CC detecting such secondary scents. This enables the construction of a scent  
 CC representation (also called a scent fingerprint or scent profile), which  
 CC may be used to re-create and edit scents. Libraries of olfactory  
 CC receptors are useful for determining the interaction pattern of a  
 CC composition with the receptors, and can be used for determining  
 CC differences in the olfactory faculties of different individuals  
 XX  
 SQ Sequence 994 BP; 201 A; 250 C; 207 G; 336 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2 21e-191 Length: 994  
 Score: 210.00 Matches: 310  
 Percent Similarity: 99.36% Conservative: 0  
 Best Local Similarity: 99.36% Mismatches: 1  
 Query Match: 67.52% Indels: 2  
 DB: 4 Gaps: 0  
 US-10-023-597-24 (1-311) x AAH31675 (1-994)  
 QY 1 MetAlAlAGluAnsSerSerSerValThrGluPheIleLeuAlAGlyLeuIleHisGln 20  
 DB 1 ATGGCTGCGAGAACCTCTCCGCGACAGAGTTTATCTCGCAGGCTTATCCACAG 60  
 QY 21 ProGlyLeuGluValProValPhePheLeuPheLeuGlyPheTyAlaValThrValAl 40  
 DB 61 CCGGAGCTCCAGGCTCCCGCTTTCTTCTGTTTCTAGGTTCTACCGGCTCAGGTTG 120  
 QY 41 GlyAsnLeuGlyLeuIleIleLeuIleGlyLeuAsnSerArgLeuHisIleProMetTy 60  
 DB 121 GGGAACTGCGGCTTGATATCTGATAGCGCTCACTCCGCTGATATCCCATGATAC 180  
 QY 61 PhePheProPheAsnLeuSerLeuValAlaPhePheSerPheSerThrIleIleProIys 80  
 DB 181 TTTTCCCTTCAACTGTCCCTCGTAGATTGTTCTCTACACACCATCTCCCAA 240  
 QY 81 MetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMetSerGln 100  
 DB 241 ATGCTGATGAGTTTCTCTCAAGAGAACATTATCTCTACAGGCGTATGAGTCAG 300  
 QY 101 PhePhePhePheCysPhePheValAlaPheSerGluSerPheIleLeuSerAlaMetValGlu 120  
 DB 301 TTTCTTTCTTCTGTTTCTTGTCTTTCTGAGTCCCTTATCTGCGCGAGTGTGTA- 359  
 QY 121 -AspAcGlyrValGlyIleCysAsnProLeuLeuTyThrIleIleMetSerProGluVal 140

DB 360 GGACCGCTACGCGGACCTGTATACCCACCTGTGTATACCATCACCTGCTCCCGAGT 419  
 QY 140 1CysLeuLeuLeuLeuLeuGlyValIleTyrglyMetGlyValPheGlyAlaValAlaHisIst 160  
 DB 420 GTGTTGCTCTCTTTACTGGGTGTCTACGGGATGGGGGTTTGGGGCTGTGCTCATAC 479  
 QY 160 rGlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyMetCysAs 180  
 DB 480 AGGAATATATAGTTTCTTCACTTTGTGACGACCAACCTGTCAATCATCATATGATGTA 539  
 QY 180 PLeuLeuProLeuLeuGluLeuSerCysAsnGlySerTyIleAsnValLeuValIlePh 200  
 DB 540 CATCTTCCCTTTTGTAGCTCTCTGCAACGGCTCTTACATATAAGTCTGTCATCTT 599  
 QY 200 eIleValAlaThrValGlyIleGlyValProIleValAlaValPheIleSerTyrglyph 220  
 DB 600 TATTGTTGTACCGTTGCGATTGGGCGGCCCATGTTGCGGTTTATCTTATAGTTT 659  
 QY 220 eIleLeuSerSerIleLeuArgValSerSerAlaGlyValArgSerIysAlaPheSerSe 240  
 DB 660 TATTCTTTCAGCATTTCTCCGCTTACTGTCTGAGGAGGAGGCTTAAACCTTCACTAG 719  
 QY 240 rCysSerSerTyIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyIle 260  
 DB 720 CTGAGGCTCTACATATTTGCAATTCTCTTTCTTTGGGTCAAGAGCTTTTACGTACT 779  
 QY 260 uLysProProSerIleLeuProLeuAspGlyValGlyValSerSerIleuPheTyThrTh 280  
 DB 780 CAAACCCCTTCCATTTTACCCCGGACGAGGGAAGTGTCTCCCTGTTCTATACAC 839  
 QY 280 rValValProMetPheAsnProLeuIleTyrSerLeuArgAsnLysAspValLysLeuAl 300  
 DB 840 TGCTGTCCCATGTTTAAACCATTAATCTACAGCTTAGAGATAAGAGATGTCAACTGCT 899  
 QY 300 aleuLysArgThrPheSerArgIleSerPheSer 311  
 DB 900 CCTGAAGAGAACCTTTTCCAGAAATTAAGCTTTTCT 933  
 RESULT 3  
 AD86616  
 ID AD86616 standard; DNA, 931 BP.  
 XX AC AD86616;  
 XX 01-JAN-2004 (first entry)  
 DT Human GPCR gene SEQ ID NO:1069.  
 DE ds; gene; human; GPCR;  
 KW guanosine triphosphate-binding protein coupled receptor; gene therapy.  
 XX Homo sapiens.  
 OS EPI270724-A2.  
 XX 02-JAN-2003.  
 PD 18-JUN-2002; 2002EP-00013517.  
 PF 18-JUN-2001; 2001JP-00246789.  
 PR (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.  
 PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.  
 XX Suwa M, Asai K, Akiyama Y, Aburatani H;  
 PI WPI; 2003-315783/31.  
 DR P-PSDB; AD86617.  
 XX New polynucleotide, useful for preparing a composition for treating a  
 PT patient in need of increased or suppressed activity or expression of the  
 PT guanosine triphosphate-binding protein coupled receptor.

XX Claim 1, SEQ ID NO 1069, 26bp, English.  
 PS  
 CC The invention relates to a novel polynucleotide encoding a guanosine  
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of  
 CC the invention may have a use in gene therapy. The polynucleotide and  
 CC polypeptide are useful for preparing a composition for treating a patient  
 CC in need of increased or suppressed activity or expression of the  
 CC guanosine triphosphate-binding protein coupled receptor. The  
 CC polynucleotide sequences shown in AD085548-AD087616 encode GPCR's of the  
 CC invention.

XX Sequence 931 BP, 202 A, 215 C, 180 G, 334 T, 0 U, 0 Other;

# Alignment Scores:

Pred. No.:	3,366-173	Length:	931
Score:	191.00	Matches:	241
Percent Similarity:	99.18%	Conservative:	1
Best Local Similarity:	99.18%	Mismatches:	2
Query Match:	61.41%	Indels:	0
DB:	9	Gaps:	0

US-10-023-597-24 (1-311) x AD086616 (1-931)

QY 70 ASPHESERPHESERTHTHTLLELPOLYSMETLEMETSERPHEVALSERARGLYS 89  
 DB 3 GATTTTAACTTCTTCAACACATCATTCCTCAAAATGCTGAGTGTGTCTCAAGAG 62  
 QY 90 AenilleleserPheThrGlyCysMetSerGlnPhePhePheCysPhePheValPhe 109  
 DB 63 AACATTATTTCTTCTCAAGGGGATGAGTCAAGTTCTTCTTCTTCTTCTTCTTCTT 122  
 QY 110 SERGLUSERPHEILLEUSERLAMETVALGLU-ASPARGLYTYVALGLYILECYSAENPR 129  
 DB 123 TCTGAGTCTTCATCTCTGCGCGATGGTGA-GGACCGGCATCGGCGCATCTTAACC 181  
 QY 129 GLEUENUTYRTHRILETHMETSERPROGLINVALCYSEUENUTLEUENUGLYVALTY 149  
 DB 182 ACTGTGTACACAGATCACCATCTCTCCCAAGTGTGTCTTCTTCTTCTTCTTCTTCTA 241  
 QY 149 TGLYMETGLYVALPHEGLYVALAVALAHLSTHRCYLAENILLEVALPHELEUTHRPHCY 169  
 DB 242 CGGAGATGGGGGTTTGGGGGCTGTGGCTCATACAGAAATATAGTGTCTTCCACCTTTTG 301  
 QY 169 SALASPHENUTLEUVALAENHISTYRMETCYASAPLLEUENUTLEUENUGLYLEUSERCY 189  
 DB 302 TGCAAGACACCTTGTCAATCACTACATGTGACATCTTCTTCTTCTTCTTCTTCTTCTG 361  
 QY 189 SANSGLYSERTYRILEASNVALLEUVALLEPHEILLEVALTHRYVALGLYILEGLYVA 209  
 DB 362 CAACGGCTCTTACATTAATGCTCTGTCATCTTATGTGTGACCGTTGGCATTTGGGGGT 421  
 QY 209 LPROLLEVALAVALPHEILLESERTYRGLYPHEILLEUSERILEUENUGLYVALSE 229  
 DB 422 GCCCATGTGTGCGTTTATCTTATGTTTATTTCTTCCAGCATCTTCCCGGTAG 481  
 QY 229 RSERLAGLUGLYARGSERLYSALAPHESESERCYSESESERTYRILEILLEALVALSE 249  
 DB 482 TTCTCTGAGGGCAGGTCTTAAGCTTCAAGTACGTCGACGCTCCATACATTAATGAGTTTC 541  
 QY 249 RLEUPHEPHEGLYSERGLYVALPHEHTYRILEULYPROFOSERILLEUENUTLEUENUG 269  
 DB 542 TCTTTTCTTGGTGGAGGAGCTTTTACGTACTCTCAACCCCTTCCATTATCCCTGGA 601  
 QY 269 PCINGLYLYSVALSERSELEUPHEHTYRTHRYVALVALPROHECPHEANPROLEUILL 289  
 DB 602 CCAAGGGAAAGTCTCTCTCTTCTTATACACATCTGTGTGCCATGTTTAAACCATTAAT 661  
 QY 289 ELYSERLEUARGANLYSAPVALYLSLEUVALLEULYSARGTHRPHESERARGLIESE 309  
 DB 662 CTACAGCCTGAGGAATAAGATGTCAAACTTCCCTGAAGAGACCTTTTCCAAATAAG 721  
 QY 309 RHESER 311

DB 722 CTTTCT 728  
 RESULT 4  
 ID ADD12748 standard; cDNA; 960 BP.  
 AC ADD12748;  
 DT 01-JAN-2004 (first entry)  
 XX  
 DE cDNA encoding novel human olfactory receptor (OR) seq id 12.  
 XX  
 KW cardiant; antidiabetic; antidiabetic; antidiabetic;  
 KW G-protein Coupled Receptor modulator; gene therapy; olfactory receptor;  
 KW G-protein Coupled Receptor X-associated disorder;  
 KW GPCR associated disorder; cardiomyopathy; atherosclerosis;  
 KW cell signal processing; metabolic pathway modulation; cancer; diabetes;  
 KW human; G-protein coupled receptor; GPCR; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003109692-A1.  
 XX  
 PD 12-JUN-2003.  
 XX  
 PE 18-DEC-2001, 2001US-00023597.  
 XX  
 PR 18-DEC-2000, 2000US-0256635P.  
 XX  
 PR 21-DEC-2000, 2000US-0257876P.  
 XX  
 PR 04-JAN-2001, 2001US-0259743P.  
 XX  
 PR 10-JAN-2001, 2001US-0260718P.  
 XX  
 PR 12-JAN-2001, 2001US-0261498P.  
 XX  
 PR 24-JAN-2001, 2001US-0263689P.  
 XX  
 PR 08-FEB-2001, 2001US-0267464P.  
 XX  
 PR 22-FEB-2001, 2001US-0271021P.  
 XX  
 PR 14-MAR-2001, 2001US-0275946P.  
 XX  
 PR 23-MAR-2001, 2001US-0278150P.  
 XX  
 PR 18-APR-2001, 2001US-0284591P.  
 XX  
 PR 23-APR-2001, 2001US-0285718P.  
 XX  
 PR 19-JUN-2001, 2001US-0299327P.  
 XX  
 PR 16-AUG-2001, 2001US-0312902P.  
 XX  
 PA (PADI/) PADIGARU M.  
 PA (KEK/) KEKUDA R.  
 PA (LIL/) LI L.  
 PA (BAL/) BALLINGER R. A.  
 PA (CAS/) CASMAN S. J.  
 PA (SPY/) SPYTEK K. A.  
 PA (BAU/) BAUMGARTNER J. C.  
 PA (BUR/) BURGESS C. E.  
 XX  
 PI Padigaru M, Kekuda R, Li L, Ballinger RA, Casman SJ, Spytek KA;  
 PI Baumgartner JC, Burgess CE;  
 DR WPI; 2003-801280/75.  
 DR P-PSDB; ADD12749.  
 XX  
 PT New isolated olfactory receptor polypeptide for treating or preventing a  
 PT G-protein Coupled Receptor associated disorder e.g. cardiomyopathy or  
 PT atherosclerosis.  
 PS  
 PS Claim 8, SEQ ID NO 23, 70pp, English.  
 XX  
 CC The invention describes an isolated olfactory receptor polypeptide (I).  
 CC (I) is used to identify an agent that binds to it by contacting (I) with  
 CC the agent and determining binding. The agent is a cellular receptor or a  
 CC downstream effector. (I) is also used to identify an agent that modulates  
 CC the expression or activity of (I) by contacting a cell expressing (I)  
 CC with the agent and determining whether the agent modulates expression or  
 CC activity of (I). (I), or nucleic acid encoding (I), is used to treat or  
 CC prevent a G-protein Coupled Receptor (GPCR)-associated disorder in a  
 CC subject, preferably a human, such as cardiomyopathy or atherosclerosis.



or a disorder related to cell signal processing or metabolic pathway modulation. (1) is used to determine the presence or predisposition to a disease associated with altered levels of (1), preferably cancer, in a mammal. A nucleic acid encoding (1) is used to determine the presence or predisposition to a disease associated with altered levels of the nucleic acid, preferably cancer, in a mammalian subject. An antibody that binds to (1) is used to treat or prevent a GPCR-associated disorder, such as diabetes or a disorder related to cell signal processing and metabolic pathway modulation. An antibody that specifically binds (1) is used to determine the presence of amount of (1) in a sample by contacting the sample with the antibody and determining the presence or amount of bound antibody. This sequence encodes a novel human olfactory receptor (OR), a G protein coupled receptor (GPCR).

Sequence 960 BP; 193 A; 247 C; 202 G; 318 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1 21e-139	Length:	960
Score:	156.00	Matches:	309
Percent Similarity:	99.36%	Conservative:	0
Best Local Similarity:	99.36%	Mismatches:	2
Query Match:	50.16%	Indels:	0
DB:	9	Gaps:	0

US-10-023-597-24 (1-311) x ADD12748 (1-960)

1 MetAlaAlaGluAsnSerSerSerValThrGluPheIleuAlaGlyLeuIleHisGln 20  
2 ATGGCTGGCCAGAACTCTCTCCGTGACAGACTTATATCCAGAGCTTAATCCACAG 61  
3 ProGlyLeuGlnValProValPhePheLeuGlyPheThrValAlaThrValVal 40  
4 CCGGAGCTCCAGATCCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121  
5 GlyAsnLeuGlyLeuIleIleuIleGlyLeuAsnSerArgLeuHisIleProMetTyr 60  
6 GGAACCTGGGCTTGATTAATCGATAGGCTCACTCTCGCATATCCCATGTC 181  
7 PhePheProPheAsnLeuSerLeuValAspPheSerPheSerThrThrIleIleProlys 80  
8 TTTTTCCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241  
9 MetLeuMetSerPheValSerArgValAsnIleIleSerPheThrGlyCysMetSerGln 100  
10 ATGCTATGATGTTTGTCTCAAGAAAGACATATTCTCTTCAAGAGGTGTATGATGAG 301  
11 PhePhePhePheCysPhePheValPheSerGlySerPheIleuSerAlaMetValGln 120  
12 TTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 361  
13 AspArgTyrValGlyIleCysAsnProLeuLeuTyrThrIlePheMetSerProGlnVal 140  
14 GACGCTACCTGGGCACTCTGTACCACTGTTGTACACGATGCTCTCCCGAGG 421  
15 CysLeuLeuLeuLeuGlyValTyrGlyMetGlyValPheGlyAlaValAlaHisThr 160  
16 TGTGTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 480  
17 GlyAsnIleValPheLeuThrPheCysAlaAspAsnLeuValAsnHisTyrMetCysAsp 180  
18 GGAATATATGATGTTTCTTCACTTGTGACAGACAACTGTGTACATACACACA--TGTGAC 538  
19 IleLeuProLeuLeuGlyLeuSerCysAsnGlySerTyrIleAsnValLeuValIlePhe 200  
20 ATCTTCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 598  
21 IleValValThrValGlyIleGlyValProIleValAlaValPheIleSerTyrGlyPhe 220  
22 ATGTGTGTGACCGTTGGCATTTGGGGTGCCCATGTGTGTGTGTGTGTGTGTGTGTGT 658  
23 IleLeuSerSerIleLeuArgValSerSerAlaGlyGlyArgSerTyrAlaPheSerSer 240  
24 ATCTTCTCCAGCATTTCTCCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 718

241 CysSerSerTyrIleIleAlaValSerLeuPhePheGlySerGlyAlaPheThrTyrLeu 260  
719 TGCAGCTCTCAATATGACGTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 778  
261 LysProProSerIleLeuProLeuAspGlnGlyValSerSerLeuPheTyrThrThr 280  
779 AAACCCCTTCAATTTTACCCCTGACAGGGAAGTGTCTCTCTCTATACCACT 838  
281 ValValProMetPheAsnProLeuIleTyrSerLeuArgAsnValAspValIleValAla 300  
839 GTGTGCTCCATGTTTAACTTATCTTACAGCTTGAGGAAATAGATGATCAAACTTGC 898  
301 LeuLysArgThrPheSerArgIleSerPheSer 311  
899 CTGAAGAGACCTTTTCCAGATATAGCTTTTCT 931

#### RESULT 5

ADCC8614  
ADCC8614 standard; DNA; 886 BP.

ADCC8614;

01-JAN-2004 (first entry)

Human GPCR gene SEQ ID NO:1067.

ds; gene; human; GPCR;

guanosine triphosphate-binding protein coupled receptor; gene therapy.

Homo sapiens.

EP1270724-A2.

02-JAN-2003.

18-JUN-2002; 2002EP-00013517.

18-JUN-2001; 2001JP-00246789.

(NAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.

(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

Suwa M, Asai K, Akiyama Y, Aburatani H;

WPI; 2003-315783/31.

P-PsDB; ADCC8615.

New polynucleotide, useful for preparing a composition for treating a

patient in need of increased or suppressed activity or expression of the

guanosine triphosphate-binding protein coupled receptor.

Claim 1; SEQ ID NO 1067; 28pp; English.

The invention relates to a novel polynucleotide encoding a guanosine

triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of

the invention may have a use in gene therapy. The polynucleotide and

polypeptide are useful for preparing a composition for treating a patient

in need of increased or suppressed activity or expression of the

CC guanosine triphosphate-binding protein coupled receptor. The

CC polynucleotide sequences shown in ADCC85548-ADCC87616 encode GPCR's of the

invention.

Sequence 886 BP; 168 A; 233 C; 176 G; 309 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	1.67e-28	Length:	886
Score:	40.00	Matches:	40
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.86%	Indels:	0
DB:	9	Gaps:	0

US-10-023-597-24 (1-311) x ADC086614 (1-886)

QY 28 PhePheLeuPheLeuGlyPheTyraIaValThrValIaGlyAsnLeuGlyLeuIleIle 47  
Db 282 TTTCTTCCTGTTTCTAGGTTTCTACCGCGTCAAGGAGGGAACCTGGCTTGATATC 341  
QY 48 LeuIleGlyLeuAsnSerArgLeuHisIlePheMetTyraPhePheProPheAsnLeuSer 67  
Db 342 CTGATAGGGCTCAACTCTCGCTGCATATCCCATGTACTTTTCCCTTCAACTGTGCC 401

RESULT 6  
AAH31668  
ID AAH31668 standard; DNA; 935 BP.

XX AAH31668;

DT 30-JUL-2001 (first entry)

DE Human olfactory receptor polynucleotide, SEQ ID NO: 241.

XX Human: olfactory receptor; OR: primary scent determination;

KW secondary scent determination; polypeptide library; odour receptor;

XX scent profile; scent fingerprint; scent representation; ds.

XX Homo sapiens.

XX WO200127158-A2.

PD 19-APR-2001.

PF 06-OCT-2000; 2000WO-US027582.

PR 08-OCT-1999; 99US-0158615P.

PR 24-FEB-2000; 2000US-0184809P.

XX (DIGI-) DIGISCENTS.

PA (YEDA) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancel D, Glusman G, Fuchs T, Yanai I,

XX WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory

PT sensation for identifying olfactory agonists and antagonists.

XX Claim 8; Page 273; 1857pp; English.

XX The present sequence is one of a number of isolated polynucleotides which

XX encode polypeptides involved in olfactory sensation. The polynucleotides

XX can be used in screening for olfactory agonists and antagonists. The

XX methods allow for the determination of primary scents and the

XX identification of the odour receptors used to detect these primary

XX scents. The methods also enable determination of secondary scents and the

XX identification of combinations of odour receptors that are involved in

XX detecting such secondary scents. This enables the construction of a scent

XX representation (also called a scent fingerprint or scent profile), which

XX may be used to re-create and edit scents. Libraries of olfactory

XX receptors are useful for determining the interaction pattern of a

XX composition with the receptors, and can be used for determining

XX differences in the olfactory faculties of different individuals

SQ Sequence 935 BP; 189 A; 238 C; 190 G; 318 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,766-28 Length: 935

Score: 40.00 Matches: 40

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.86% Indels: 0

DB: 4 Gaps: 0

US-10-023-597-24 (1-311) x AAH31668 (1-935)

QY 28 PhePheLeuPheLeuGlyPheTyraIaValThrValIaGlyAsnLeuGlyLeuIleIle 47

Db 82 TTTCTTCCTGTTTCTAGGTTTCTACCGCGTCAAGGAGGGAACCTGGCTTGATATC 141

QY 48 LeuIleGlyLeuAsnSerArgLeuHisIlePheMetTyraPhePheProPheAsnLeuSer 67

Db 142 CTGATAGGGCTCAACTCTCGCTGCATATCCCATGTACTTTTCCCTTCAACTGTGCC 201

RESULT 7  
AAH84099  
ID AAH84099 standard; DNA; 487 BP.

XX AAH84099;

DT 25-SEP-2001 (first entry)

DE Mouse olfactory receptor encoding gene 6.

XX Olfactory receptor; primate; mouse; human; food processing industry;

KW aromas; perfumery; toxic substance; ds.

XX Mus musculus domesticus.

XX WO200146262-A2.

PD 28-JUN-2001.

PF 22-DEC-2000; 2000WO-IB002017.

PR 22-DEC-1999; 99US-0171746P.

PR 21-DEC-2000; 2000US-00747155.

XX (CNRS) CNRS CENT NAT RECH SCI.

XX Rouquier S, Giorgi D;

XX WPI; 2001-381911/40.

XX P-PDB; AAG98384.

XX Nucleic acids encoding primate and murine olfactory receptors, useful for

PT analysis odors e.g. in food processing and perfumery.

XX Claim 1; Page 415-416; 482pp; English.

XX The invention relates to olfactory receptors (AAG98432-AAG98609) and the

XX genes encoding them (AAH83879-AAH84131) including pseudogenes of 10

XX CC primate species, mouse and human. The nucleic acids and receptors may be

XX used in the food processing industry (e.g. for the detection of aromas,

XX CC quality control and sample analysis), in perfumery (e.g. for the analysis

XX or comparison of perfumes) and in the environment (e.g. for the detection

XX of toxic substances and/or trapping of odours)

SQ Sequence 487 BP; 86 A; 120 C; 105 G; 176 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,786-22 Length: 487

Score: 33.00 Matches: 33

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.61% Indels: 0

DB: 4 Gaps: 0

US-10-023-597-24 (1-311) x AAH84099 (1-487)

QY 172 AsnLeuValAsnHisTyraValMetCysAspIleLeuProLeuLeuGlyLeuSerCysAsnGly 191

Db 146 AACCTGTGAATCACTACATGTGTAACATCTCCCTTCTGAGCTCTCCGCAATGAC 205

QY 192 SerTyraLeuAsnValLeuValIlePheIleValValThr 204

Db 206 TCTTACATAAATGTTCTGTGTCATCTTATGTGTGAC 244

RESULT 8

AAH84021  
ID AAH84021 standard; DNA; 485 BP.  
XX  
XX  
AC AAH84021,  
XX  
XX 06-AUG-2003 (revised)  
DT 25-SEP-2001 (first entry)  
XX  
XX Bulemar fulvus olfactory receptor encoding gene 5.  
DE  
XX  
KM Olfactory receptor; primate; mouse; human; food processing industry;  
KM atomas; perfume; toxic substance; ds.  
XX  
OS Bulemar fulvus.  
XX  
PN WO200146262-A2.  
XX  
XX 28-JUN-2001.  
XX  
XX 22-DEC-2000; 2000WO-1B002017.  
PF  
XX 22-DEC-1999; 99US-0171746P.  
PR 21-DEC-2000; 2000US-00747155.  
XX  
XX (CNRS ) CNRS CENT NAT RECH SCI.  
PA  
XX  
PI Rouquier S, Giorgi D;  
XX  
XX WPI; 2001-381911/40.  
DR  
XX P-PSDB; AAG8524.  
XX  
XX Nucleic acids encoding primate and murine olfactory receptors, useful for  
PT analysis odors e.g. in food processing and perfumery.  
XX  
PS Claim 1; Page 280-281; 482pp; English.  
XX  
XX The invention relates to olfactory receptors (AAG98432-AAG98609) and the  
CC genes encoding them (AAH83879-AAH84131) including pseudogenes of 10  
CC primate species, mouse and human. The nucleic acids and receptors may be  
CC used in the food processing industry (e.g. for the detection of aromas,  
CC quality control and sample analysis), in perfumery (e.g. for the analysis  
CC or comparison of perfumes) and in the environment (e.g. for the detection  
CC of toxic substances and/or trapping of odours). (Updated on 06-AUG-2003  
CC to correct OS field.)  
XX  
XX  
SQ Sequence 485 BP; 90 A; 116 C; 106 G; 173 T; 0 U; 0 Other;  
XX  
XX  
Alignment Scores:  
Pred. No.: 1,37e-09 Length: 485  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
Gaps: 0  
DB: 4  
XX  
US-10-023-597-24 (1-311) x AAH84021 (1-485)  
OY 171 AspAsnLeuValAsnHisTyrMetCysAspIleLeuProLeuLeuGluLeuSerCysAsn 190  
DB 142 GACAACTGTGCAATCACTAATATGTGACATCTCTCTTGAGCTCTCTGCAAC 201  
XX  
XX  
RESULT 9  
AAH32421  
ID AAH32421 standard; DNA; 930 BP.  
XX  
XX  
AC AAH32421,  
XX  
XX 30-JUL-2001 (first entry)  
DT  
XX  
XX Human olfactory receptor polynucleotide, SEQ ID NO: 994.  
DE  
XX  
XX Human; olfactory receptor; OR; primary scent determination;  
KM secondary scent determination; polypeptide library; odour receptor;  
XX

KM scent profile; scent fingerprint; scent representation; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200127158-A2.  
PN  
XX  
XX 19-APR-2001.  
PD  
XX  
XX 06-OCT-2000; 2000WO-US027582.  
PF  
XX  
XX 08-OCT-1999; 99US-0158615P.  
PR 24-FEB-2000; 2000US-0184809P.  
XX  
XX (DIGI-) DIGISCENTS.  
PA (YEDA ) YEDA RES & DEV CO LTD.  
XX  
XX  
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX  
XX WPI; 2001-290713/30.  
DR  
XX  
XX New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists.  
XX  
XX  
PS Claim 8; Page 577; 1857pp; English.  
XX  
XX The present sequence is one of a number of isolated polynucleotides which  
CC encode polypeptides involved in olfactory sensation. The polynucleotides  
CC can be used in screening for olfactory agonists and antagonists. The  
CC methods allow for the determination of primary scents and the  
CC identification of the odour receptors used to detect these primary  
CC scents. The methods also enable determination of secondary scents and the  
CC identification of combinations of odour receptors that are involved in  
CC detecting such secondary scents. This enables the construction of a scent  
CC representation (also called a scent fingerprint or scent profile), which  
CC may be used to re-create and edit scents. Libraries of olfactory  
CC receptors are useful for determining the interaction pattern of a  
CC composition with the receptors, and can be used for determining  
CC differences in the olfactory faculties of different individuals  
XX  
XX  
SQ Sequence 930 BP; 190 A; 251 C; 188 G; 301 T; 0 U; 0 Other;  
XX  
XX  
Alignment Scores:  
Pred. No.: 2,58e-09 Length: 930  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
Gaps: 0  
DB: 4  
XX  
US-10-023-597-24 (1-311) x AAH32421 (1-930)  
OY 79 ProlysMetLeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet 98  
DB 232 CCGAAATGCTGATGAGTGTGTTGTTCTCAAGAGAACATCAATTCCTTCACAGGGTGTATG 291  
XX  
XX  
RESULT 10  
AAS15910  
ID AAS15910 standard; cDNA; 933 BP.  
XX  
XX  
AC AAS15910;  
XX  
XX 25-JAN-2002 (first entry)  
DT  
XX  
XX DNA encoding G-protein coupled receptor (GCRC) #14.  
DE  
XX  
XX G-protein coupled receptor; GCRC; vaccine; gene therapy;  
KM cell proliferation disorder; cancer; arteriosclerosis;  
KM neurological disorder; epilepsy; stroke; cardiovascular disorder;  
KM hypertension; ischaemic heart disease; gastrointestinal disorder;  
KM anorexia; peptic ulcer; autoimmune disorder; inflammatory disorder;  
KM diabetes; osteoporosis; psoriasis; metabolic disorder; obesity;  
KM schizophrenic disorder; neuroskeletal disorder; ss.  
XX

OS Homo sapiens.  
 XX Key Location/Qualifiers  
 FH 1. 933  
 FT CDS  
 FT /tag= a  
 FT /product= "GCREC 14"  
 FT /note= "G-protein coupled receptor 14"  
 PN W020016742-A2.  
 PD 13-SEP-2001.  
 PF 01-MAR-2001; 2001WO-US006814.  
 XX 03-MAR-2000; 2000US-0186854P.  
 XX 10-MAR-2000; 2000US-0188384P.  
 PR 17-MAR-2000; 2000US-0190453P.  
 PR 20-MAR-2000; 2000US-0190730P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 XX Lal P, Tang YT, Patterson C, Yao MG, Shih LL, Tribouley CM,  
 PI Lu DM, Yue H, Khan FA, Policky JL, Au-Young J, Yang J, Harland L,  
 PI Walsh RT, Lo TP, Borowsky NL,  
 XX WPI. 2001-656776/75.  
 DR P-FSDB; AAU10313.  
 XX  
 PT Novel G-protein coupled receptor polypeptides, for treating and  
 PT preventing autoimmune/inflammatory disorders, neurological disorders,  
 PT cell proliferative disorders, cardiovascular disorders and viral  
 PT infections.  
 XX  
 PS Claim 5, Page 138; 141pp; English.  
 XX  
 CC The invention describes a novel isolated polypeptide, selected from a  
 CC group of 21 G-protein coupled receptor polypeptides (GCREC) and useful in  
 CC vaccines and gene therapy. The polypeptide (I) is useful for screening  
 CC for agonist or antagonist of (I), compounds specifically binding to (I),  
 CC or compounds that modulate the activity of (I). The polynucleotide  
 CC encoding (I) is useful for screening a compound for effectiveness in  
 CC altering expression of a target polynucleotide comprising (II), by  
 CC exposing a sample comprising the target polynucleotide to a compound,  
 CC detecting altered expression of the target polynucleotide, and comparing  
 CC the expression of the target polynucleotide in the presence of varying  
 CC amounts of compound and in the absence of the compound. (I) and (II) are  
 CC useful for diagnosis, treatment and prevention of cell proliferative  
 CC disorders (e.g. cancers, arteriosclerosis, atherosclerosis), neurological  
 CC disorders (e.g. epilepsy, stroke, schizophrenia disorders and  
 CC neuromuscular disorders), cardiovascular disorders (e.g. hypertension,  
 CC ischemic heart disease), gastrointestinal disorders (e.g. diabetes mellitus,  
 CC peptic ulcer), autoimmune/inflammatory disorders (e.g. diabetes mellitus,  
 CC osteoporosis, psoriasis), and metabolic disorders such as obesity.  
 CC Furthermore, the polynucleotide is useful, as primers for detecting  
 CC single nucleotide polymorphisms, as elements in microarray, to monitor or  
 CC measure protein-protein interactions, drug-target interactions, and gene  
 CC expression profiles; to generate a transcript image of a tissue or cell  
 CC type, and to generate hybridisation probes useful in mapping the  
 CC naturally occurring genomic sequence. This sequence encodes G-protein  
 CC coupled receptor 14, one of 21 GCREC proteins described in the method of  
 CC the invention  
 XX  
 SQ Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 2.58e-09 Length: 933  
 Score: 20.00 Matches: 20  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.43% Indels: 0  
 DB: Gaps: 0  
 US-10-023-597-24 (1-311) X AAS15910 (1-933)

Qy 79 ProlysmetLeuMetSerPheValSerArglyAsaenilelleSerPheThrGlyCysMet 98  
 Db 232 CCAGAAATGCTGATGATTTGCTCAAGCAAGAACATCATTTCTTCACAGGGGTATG 291  
 RESULT 11  
 ID AAS42233  
 ID AAS42233 standard; cDNA, 933 BP.  
 AC AAS42233;  
 XX  
 AC AAS42233;  
 XX  
 DT 18-DEC-2001 (first entry)  
 XX  
 DE Human cDNA encoding olfactory receptor AOLF26.  
 XX  
 XX Human; olfactory receptor; G protein-coupled receptor; GPCR; odorant;  
 XX sm; food additive; cosmetic; fragrance; pharmaceutical additive.  
 OS Homo sapiens.  
 XX  
 XX W0200168805-A2.  
 XX  
 PD 20-SEP-2001.  
 XX  
 PF 13-MAR-2001; 2001WO-US007771.  
 XX  
 PR 13-MAR-2000; 2000US-0188914P.  
 PR 24-MAR-2000; 2000US-0192033P.  
 PR 12-APR-2000; 2000US-0198474P.  
 PR 24-APR-2000; 2000US-0199335P.  
 PR 26-MAY-2000; 2000US-0207702P.  
 PR 23-JUN-2000; 2000US-0213849P.  
 PR 16-AUG-2000; 2000US-0226534P.  
 PR 07-SEP-2000; 2000US-0230732P.  
 PR 07-FEB-2001; 2001US-0266662P.  
 XX  
 PA (SENO-) SENOMTX INC.  
 XX  
 PI Zozulya S;  
 XX  
 DR WPI. 2001-570867/64.  
 DR P-FSDB; AAU24540.  
 XX  
 PT Nucleic acids encoding human olfactory G protein-coupled receptors,  
 PT useful for screening for compounds involved in olfactory sensation, where  
 PT the compounds can be used in the food, pharmaceutical and cosmetic  
 PT industries to customize odors.  
 XX  
 PS Claim 1; Page 94; 319pp; English.  
 XX  
 CC The invention relates to nucleic acids encoding human olfactory  
 CC receptors, OR, (a G protein-coupled receptor, GPCR). The OR's  
 CC specifically recognise molecules, odourants, that elicit specific  
 CC olfactory sensation. The human olfactory receptors and polynucleotides  
 CC encoding them are useful for screening a library of chemical compounds  
 CC for compounds that are involved in olfactory sensation. Modulators of  
 CC their activity are useful for pharmacological and genetic modulation of  
 CC olfactory signalling pathways. Therefore, they can be used in the food,  
 CC pharmaceutical and cosmetic industries to customise odours and  
 CC fragrances. The present sequence encodes a human olfactory receptor of  
 CC the invention  
 XX  
 SQ Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;  
 XX  
 Alignment Scores:  
 Pred. No.: 2.58e-09 Length: 933  
 Score: 20.00 Matches: 20  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.43% Indels: 0  
 DB: Gaps: 0  
 US-10-023-597-24 (1-311) X AAS42233 (1-933)

QY 79 ProlyMetLeuMetSerPheValSerArglyAsnIleIleSerPheThrGlyCysMet 98  
DB 232 CCMAAATGCTGATGATTTCCTCAGGAGAACATCATTTCTTCACAGGCTATG 291

RESULT 12  
ABZ43067  
ID ABZ43067 standard; DNA; 933 BP.  
XX  
XX ABZ43067;  
AC  
XX  
XX 06-MAR-2003 (first entry)  
DT  
XX  
XX Human GPCR polynucleotide SEQ ID NO 395.  
DE  
XX  
XX Human; GPCR; G protein coupled receptor; signal transduction; olfactory;  
KM drug development; gustatory; taste; fragrance; gene; ds.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200216548-A2.  
PN  
XX  
XX 28-FEB-2002.  
PD  
XX  
XX 30-JUL-2001; 2001WO-IB001446.  
PF  
XX  
XX 04-AUG-2000; 2000JP-00237818.  
PR  
XX  
XX 13-FEB-2001; 2001JP-00034434.  
PA  
XX  
XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.  
PI  
XX  
XX Haga T, Takeda S, Mitaku S;  
DR  
XX  
XX WPI; 2002-304118/34.  
DR  
XX  
XX P-PSDB; ABP95793.  
PT  
XX  
XX Database global search for G protein-coupled receptors, proteins and  
PT encoded genes for studying in vivo signal transduction mechanism and  
PT identifying targets for drug development.

PS  
XX  
XX Claim 9; SEQ ID NO 395; 97pp + Sequence Listing; Japanese.  
PS  
XX  
XX The invention relates to a method for screening G protein-coupled  
CC receptor (GPCR) genes (ABZ42870-ABZ43216) and/or GPCR proteins (ABP95596-  
CC ABP95942) by extracting open-reading frames containing 6-8 transmembrane  
CC domains with 250-1000 amino acid residues to give a gene homologous with  
CC a known GPCR gene. The receptor proteins and encoded genes are useful for  
CC studying in vivo signal transduction mechanism and identifying targets  
CC for drug development e.g. based on olfactory and gustatory receptors in  
CC form of agonists and antagonists by screening intrinsic and extrinsic  
CC ligands as bitter taste inhibitors, taste enhancers and fragrance  
CC improves. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pat\_sequences  
XX  
XX  
SQ Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.58e-09 Length: 933  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 6 Gaps: 0

US-10-023-597-24 (1-311) x ABZ43067 (1-933)

QY 79 ProlyMetLeuMetSerPheValSerArglyAsnIleIleSerPheThrGlyCysMet 98  
DB 232 CCMAAATGCTGATGATTTCCTCAGGAGAACATCATTTCTTCACAGGCTATG 291

RESULT 13  
ABK68455

ID ABK68455 standard; DNA; 933 BP.  
XX  
XX  
XX ABK68455;  
AC  
XX  
XX 02-JUL-2002 (first entry)  
DT  
XX  
XX  
DE Human DNA for olfactory and pheromone G protein-coupled receptor #55.  
XX  
XX  
XX Human; ds; gene; olfactory and pheromone G protein coupled receptor;  
KM GPCR; tranquilizer; antidepressant; neuroleptic; endocrine; anabolic;  
KM anorectic; taste; fragrance; food additive; cosmetic; cell migration;  
KM sterility; psychotic disorder; neurological disorder; anxiety;  
KM schizophrenia; manic depression; depression; axonal growth;  
KM menstrual cycle; appetite sexual motivation; sexual attraction;  
KM aggression.  
XX  
XX  
XX Homo sapiens.  
OS  
XX  
XX  
XX WO200224726-A2.  
FN  
XX  
XX  
XX 28-MAR-2002.  
PD  
XX  
XX  
XX 21-SEP-2001; 2001WO-BE000162.  
PF  
XX  
XX  
XX 22-SEP-2000; 2000EP-00870211.  
PR  
XX  
XX  
XX (CHEM-) CHEMCOM SA.  
PA  
XX  
XX  
XX Velthen A;  
PI  
XX  
XX  
XX WPI; 2002-330013/36.  
DR  
XX  
XX  
XX P-PSDB; AAU95568.  
DR  
XX  
XX  
XX Novel pheromone G-protein coupled receptor and receptor-derived agonists,  
PT antagonists or inhibitors useful in food or cosmetic products or in the  
PT treatment or prevention of neurological disorders such as anxiety and  
PT schizophrenia.

PS  
XX  
XX Disclosure; Page 178-179; 833pp; English.  
PS  
XX  
XX  
XX The invention relates to olfactory and pheromone G-protein coupled  
CC receptor (GPCR) or a protein 95% identical to the GPCR, a specific active  
CC portion and its encoding polynucleotide. Also included are an agonist,  
CC antagonist or inhibitor of the GPCR or the polynucleotide, a vector  
CC comprising the polynucleotide, a cell transformed by the vector, a non-  
CC human mammal comprising a partial or total deletion of the polynucleotide  
CC encoding the receptor and screening (detection and possibly, recovering)  
CC of compounds which are known or not known to be agonist, antagonists or  
CC inhibitors of natural compounds to the GPCR. The receptor-derived  
CC agonists, antagonists, inhibitors or compounds are used as an  
CC improvement, elimination or substitution of an existing taste and/or a  
CC fragrance of (or in) the food and/or cosmetic products. They can also be  
CC used in the preparation of medicament in the treatment and/or prevention  
CC of a mammalian disorder, such as cell migration, sterility, psychotic and  
CC neurological disorders, including anxiety, schizophrenia, manic  
CC depression, depression, for promoting axonal growth, nerve cell  
CC connection and nerve regeneration for modulating male and female  
CC endocrine functions, hormone production and the menstrual cycle, for the  
CC prevention or the treatment by stimulation of several mammalian  
CC behaviours, such as stimulation or suppression of appetite, sexual  
CC motivation, sexual attraction, aggression and for promoting or  
CC suppressing chemical communication between organisms. The present  
CC sequence is a human DNA encoding an olfactory and pheromone GPCR  
XX  
XX  
SQ Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 2.58e-09 Length: 933  
Score: 20.00 Matches: 20  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.43% Indels: 0  
DB: 6 Gaps: 0

US-10-023-597-24 (1-311) x ABK68455 (1-933)

Qy 79 ProlysmetLeuMetSerPheValSerArgIysAsnIleIleSerPheThrGlyCysMet 98  
 ID 232 CCCAAATGCTGATGATTGTTGCTCAGAGAAAGAACATCATCTTCTTCACAGGGGTGATG 291

RESULT 14  
 ABK37521  
 ID ABK37521 standard; cDNA; 933 BP.

AC ABK37521;

DT 08-MAY-2002 (first entry)

DE DNA encoding G-coupled olfactory receptor #23.

KW Human; olfactory G-coupled receptor; sensory perception of odourant;  
 KM odour composition; taste composition; gene; ss.

OS Homo sapiens.

PN WO200198526-A2.

PD 27-DEC-2001.

PF 22-JUN-2001; 2001WO-US020122.

PR 22-JUN-2000; 2000US-0213812P.

PR 13-MAR-2001; 2001US-00804291.

PA (SENO-) SENOMYX INC.

PI Zozulya S, Stryer L;

DR WPI, 2002-083330/11.

DR P-PSDB; AAC85162.

PT Representing sensory perception of one or more odourants for the  
 PT identification and design of tastes and odors comprises providing a  
 PT representative group of n olfactory receptors.

XX Example; Page 68; 182pp; English.

CC The invention relates to a method of representing sensory perception of  
 CC one or more odourants. The method comprises: (a) providing a  
 CC representative class of n olfactory receptors or ligand binding domains of  
 CC (LBD) of these receptors; (b) measuring values X1 to Xn representative of  
 CC at least one activity of one or more odourants selected from: (i) binding  
 CC one or more odourants to the LBD of at least one of the n olfactory  
 CC receptors; (ii) activating at least one of the n olfactory receptors with  
 CC the one or more odourants; and (iii) blocking at least one of the n  
 CC olfactory receptors with the one or more odourants; and (c) generating a  
 CC representation of sensory perception from the values X1 to Xn. The  
 CC design and formulation of odour and taste compositions. ABK37499-ABK37754  
 CC and ABK37918-ABK37921 represent human olfactory G-coupled receptor coding  
 CC sequences and related PCR primers of the invention

CC Sequence 933 BP; 192 A; 251 C; 188 G; 302 T; 0 U; 0 Other;

SO Alignment Scores:

Prod. No.: 2,58e-09 Length: 933  
 Score: 20.00 Matches: 20  
 Percent Similarity: 100.00% Conservatve: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.43% Indels: 0  
 DB: Gaps: 0

US-10-023-597-24 (1-311) x ABK37521 (1-933)

Qy 79 ProlysmetLeuMetSerPheValSerArgIysAsnIleIleSerPheThrGlyCysMet 98  
 ID 232 CCCAAATGCTGATGATTGTTGCTCAGAGAAAGAACATCATCTTCTTCACAGGGGTGATG 291

DB 232 CCCAAATGCTGATGATTGTTGCTCAGAGAAAGAACATCATCTTCTTCACAGGGGTGATG 291

RESULT 15

ID ABT05655  
 ID ABT05655 standard; DNA; 953 BP.

AC ABT05655;

DT 16-OCT-2002 (first entry)

DE GPCR 4 protein encoding DNA SEQ ID No 13.

KW Antidiabetic; cytosolic; anorectic; nootropic; neuroprotective; GPCR;  
 KW antiparkinsonian; cardiac; antihypertensive; immunosuppressive;  
 KW hypotensive; haemostatic; antifertility; antiaesthetic; antiinflammatory;  
 KW anti-HIV; G-protein coupled receptor X; cardiomyopathy; atherosclerosis;  
 KW diabetes; metabolic disorder; diabetes; obesity; infectious disease;  
 KW anorexia; neurodegenerative disorder; Alzheimer's disease; dyslipidaemia;  
 KW Parkinson's disorder; immune disorder; haematopoietic disorder; obesity;  
 KW metabolic syndrome X; wasting disorder; cancer; hypertension; neoplasm;  
 KW congenital heart defect; aortic stenosis; subaortic stenosis; lymphoma;  
 KW transplantation; adrenoleukodystrophy; congenital adrenal hyperplasia;  
 KW prostate cancer; adenocarcinoma; uterus cancer; fertility; haemophilia;  
 KW hypercoagulation; idiopathic thrombocytopenic purpura; immunodeficiency;  
 KW graft versus host disease; AIDS; bronchial asthma; Crohn's disease;  
 KW multiple sclerosis; chromosomal mapping; tissue typing; forensic biology;  
 KW gene therapy; transgenic animal; gene; ds.

OS Unidentified.

PN WO200246229-A2.

PD 13-JUN-2002.

PF 05-DEC-2001; 2001WO-US046530.

PR 05-DEC-2000; 2000US-0251459P.

PR 29-DEC-2000; 2000US-0259007P.

PR 04-DEC-2001; 2001US-00005041.

PA (CURA-) CURAGEN CORP.

PI Casman SJ, Padigaru M, Burgess CE, Shinkets RA, Spytek KA;

PI Gilbert JA, Nayotte JE, Baumgartner JC, Mishra V, Vernet CAM;

PI Dickinson KS, Ballinger RA, Wolenc AR;

DR WPI; 2002-537559/57.

DR P-PSDB; AB04724.

PT Novel isolated G-protein coupled receptor polypeptide, designated NOVX,  
 PT useful for treating or preventing n human receptor-associated disorders  
 PT e.g. cardiomyopathy, atherosclerosis or diabetes.

XX Claim 9; Page 33; 264pp; English.

CC The invention relates to a novel isolated G-protein coupled receptor X  
 CC (GPCRX) polypeptide. The isolated protein, its encoding polynucleotide  
 CC and the antibody of the isolated protein is useful for treating or  
 CC preventing a GPCR-associated disorder in a subject, preferably human,  
 CC where the disorder is cardiomyopathy, atherosclerosis or diabetes. The  
 CC isolated protein, its encoding polynucleotide and the antibody of the  
 CC isolated protein is useful for treating or preventing metabolic  
 CC disorders, diabetes, obesity, infectious disease, anorexia,  
 CC neurodegenerative disorder, Alzheimer's disease, Parkinson's disorder,  
 CC immune disorders, haematopoietic disorders, and various dyslipidaemias,  
 CC metabolic disturbances associated with chronic diseases, and cancer. The  
 CC wasting disorders associated with chronic diseases, and cancer. The  
 CC disorders also include cardiomyopathy, atherosclerosis, hypertension,  
 CC congenital heart defects, aortic stenosis, subaortic stenosis,  
 CC transplantation, adrenoleukodystrophy, congenital adrenal hyperplasia,  
 CC prostate cancer, neoplasm, adenocarcinoma, lymphoma, uterus cancer,  
 CC fertility, haemophilia, hypercoagulation, idiopathic thrombocytopenic  
 CC purpura, immunodeficiencies, graft versus host disease, AIDS, bronchial

CC asthma, Crohn's disease, and multiple sclerosis. The isolated protein,  
 CC its encoding polynucleotide and the antibody of the isolated protein is  
 CC useful in screening assays, detection assays (e.g., chromosomal mapping,  
 CC tissue typing, forensic biology). The isolated polynucleotide is useful  
 CC in gene therapy, to express the isolated protein, to detect GPCR mRNA or  
 CC a genetic lesion in a GPCR gene, and to modulate GPCR activity. The  
 CC cell of the invention is useful for producing non-human transgenic  
 CC animals. This polynucleotide sequence represents DNA encoding a GPCR  
 CC protein of the invention  
 XX

SQ Sequence 953 BP; 202 A; 256 C; 193 G; 302 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,64e-09	Length:	953
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.43%	Indels:	0
DB:	6	Gaps:	0

US-10-023-597-24 (1-311) X APT05655 (1-953)

OY	79	ProLysMetIeuMetSerPheValSerArgLysAsnIleIleSerPheThrGlyCysMet	98
DB	247	CCCAAAATGCTGATGAGTTTGTCTCAAGAGAAACATCATTTCTTCAAGGGTGATG	306

Search completed: September 30, 2004, 14:02:34  
 Job time : 451 secs

*This Page Blank (uspto)*